

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 01:15:17 ; Search time 2583 Seconds
(without alignments)
9541.130 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt.....tgaagaaaacggcccttga 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*

28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query				ID	Description
	No.	Score	Match Length	DB		
	1	260.4	84.9	1552	11 AK076867	AK076867 Mus muscu
	2	860.4	84.9	2245	11 AK030474	AK030474 Mus muscu
	3	260.4	84.9	3039	11 AK053642	AK053642 Mus muscu
	4	358.8	84.7	1377	11 AK076758	AK076758 Mus muscu
	5	844.8	83.3	1449	11 AK013205	AK013205 Mus muscu
	6	770.6	76.0	822	9 AU125107	AU125107 AU125107
	7	750.8	74.0	1201	13 BX393735	BX393735 BX393735
	8	709.4	70.0	1379	11 AK005323	AK005323 Mus muscu
	9	571.8	66.3	784	10 BG218735	BG218735 RST38476
	10	622.8	61.4	1281	11 AK013161	AK013161 Mus muscu
	11	614	60.6	951	13 BU116522	BU116522 603139786
	12	594.2	58.6	982	13 BQ669953	BQ669953 AGENCOURT
	13	585.2	57.7	934	13 BU518807	BU518807 AGENCOURT
	14	579.2	57.1	713	14 CD354831	CD354831 UI-M-GMO-
	15	576	57.0	742	2 HSM073180	Bx483012 Homo sapi
	16	565	55.7	946	14 CA973078	CA973078 AGENCOURT
	17	553.4	54.6	958	13 BU514920	BU514920 AGENCOURT
	18	550.2	54.3	563	9 AA278473	AA278473 zs81h12.r
	19	542.6	53.5	930	14 CA982682	CA982682 AGENCOURT
	20	536.2	52.9	732	14 CD103801	CD103801 AGENCOURT
	21	533.2	52.6	1186	10 BF159587	BF159587 601769084
	22	521.4	51.4	721	9 AW242839	AW242839 xn26f05.x
	23	517.8	51.1	985	13 BU914186	BU914186 AGENCOURT
	24	500.4	49.3	1060	13 BQ899617	BQ899617 AGENCOURT
	25	499.6	49.3	817	14 CB559060	CB559060 AGENCOURT
	26	491.8	48.5	893	13 BU367571	BU367571 603787711
	27	486.6	48.0	842	13 BQ940647	BQ940647 AGENCOURT
	28	473.6	46.7	854	14 CB200465	CB200465 AGENCOURT
	29	461	45.5	634	12 BQ021348	BQ021348 UI-H-DH1-
	30	460.8	45.4	933	14 CA983082	CA983082 AGENCOURT
	31	457.6	45.1	575	12 BI499153	BI499153 ie27c05.y
	32	452	44.6	863	13 BU905264	BU905264 AGENCOURT
	33	451.4	44.5	927	12 BI655225	BI655225 603284039
	34	451.2	44.5	876	13 BQ880200	BQ880200 AGENCOURT
	35	445.2	43.9	955	13 BU152376	BU152376 AGENCOURT
	36	445	43.9	534	9 AI645170	AI645170 mt82a02.y
	37	443.8	43.8	833	9 AL961440	AL961440 AL961440
	38	442.8	43.7	853	14 CB564338	CB564338 AGENCOURT
	39	441.6	43.6	697	12 BI828691	BI828691 603074707
	40	435.8	43.0	1046	14 CB235478	CB235478 AGENCOURT
	41	435.2	42.9	440	9 AA669484	AA669484 af74g07.r
	42	434.2	42.8	662	13 BU631151	BU631151 UI-H-FE1-
	43	433	42.7	961	13 BU129770	BU129770 603117261
	44	430	42.4	752	14 CB524885	CB524885 UI-M-FY0-
	45	428	42.2	985	14 CA970822	CA970822 AGENCOURT

ALIGNMENTS

RESULT 1

AK076867

LCCUS AK076867 1552 bp mRNA linear HTC 07-DEC-2002

DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930520C08 product:MO25-LIKE PROTEIN homolog [Homo sapiens], full insert sequence.

ACCESSION AK076867

VERSION AK076867.1 GI:26345723

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzairelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1552)
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Ncmura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES Location/Qualifiers
source 1. .1552
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:493C520C03"
/db_xref="MGI:1894876"
/db_xref="taxon:10090"
/clone="4930520C03"
/sex="male"

CDS

```
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
316..1320
/note="unnamed protein product; MO25-LIKE PROTEIN homolog
[Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
98.2%ID, 100%length, match=1002)
putative"
/codon_start=1
/protein_id="BAC36513.1"
/db_xref="GI:26345724"
/db_xref="MGI:1914081"
/translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ
AMKEILCGTNDKEPPTTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL
RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN
YVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
"
```

```
polyA_signal 1539..1544
              /note="putative"
```

```
polyA_site 1552
            /note="putative"
```

```
BASE COUNT 490 a 320 c 341 g 401 t
ORIGIN
```

```
Query Match 84.9%; Score 860.4; DB 11; Length 1552;
Best Local Similarity 90.5%; Pred. No. 5.8e-176;
Matches 918; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
   |||
DB 307 ATGAAAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 366
   |||

QY 61 CTGAAAGACAAATTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
   |||
DB 367 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 426
   |||

QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
   || |||
DB 427 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 486
   || |||

QY 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
   ||| |||
DB 437 CCCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGTTGCTGGTG 546
   ||| |||

QY 241 AACTGATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 300
   ||| |||
DB 547 AACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAAGATGTGACCCAGATA 606
   ||| |||

QY 301 TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
   || |||
DB 607 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCTACTGTGAGTACATCAGT 666
   || |||

QY 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
   ||| |||
DB 667 TCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 726
   ||| |||
```


Qy 421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
 || |||||
 Db 727 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 786
 Qy 431 TTTTCTAATCAATTGAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||||
 Db 787 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 846
 Qy 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 |||||
 Db 847 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 906
 Qy 501 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 |||||
 Db 207 TTCTTAGAACAAAATTATGACACTATTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 966
 Qy 561 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 || |||||
 Db 567 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1026
 Qy 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
 || || |||||
 Db 1027 AATTTCACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1086
 Qy 701 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 || |||||
 Db 1087 CTGCTTCGAGACAAAAGTCCCAACATCCAATTGGAAGCCTTCCATGTCTTTAAGGTGTTT 1146
 Qy 941 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAA 900
 |||||
 Db 1147 GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTAAAAAATCAGCCCAAA 1206
 Qy 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 |||||
 Db 1207 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1266
 Qy 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 |||||
 Db 1267 GAGAAGAACTACCTGATTAAACAGATTGAGACTTGAAGAAAGCAGCCCCCTGA 1320

RESUME

AK030474

LOCUS AK030474 2245 bp mRNA linear HTC 05-DEC-2002

DEFINITION Mus musculus adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:5330416K15 product:MO25-LIKE PROTEIN homolog [Homo sapiens], full insert sequence.

ACCESSION AK030474

VERSION AK030474.1 GI:26326468

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staebli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.

TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6 (bases 1 to 2245)
 AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,

Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/.

FEATURES

SOURCE

Location/Qualifiers

1. .2245
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:5330416K15"
/db_xref="taxon:10090"
/clone="5330416K15"
/sex="male"
/tissue_type="pituitary gland"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

CDS

300. .1313
/note="unnamed protein product; MO25-LIKE PROTEIN homolog [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY, 98.2%ID, 100%length, match=1002) putative"
/codon_start=1
/protein_id="BAC26978.1"
/db_xref="GI:26326469"
/translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK SLQAMKEILCGTNDKEPPTTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL FSNQFRDFFKYVELSTFDIASDAFATFKDLITRHKVLVADFLEQNYDTIFEDYEKLLQ SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKEPTDDEQFADEKNYLIKQIRDLKK AAP"

BASE COUNT
ORIGIN

667 a 480 c 517 g 581 t

Query Match 84.9%; Score 860.4; DB 11; Length 2245;
Best Local Similarity 90.5%; Pred. No. 6.2e-176;
Matches 918; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
QY      1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
      |||
Db      300 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 359

QY      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
      |||
Db      360 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 419

QY     121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
      |||
Db     420 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGACACAAGGAG 479

QY     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG 240
      |||
Db     480 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 539

QY     241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
      |||
Db     540 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 599

QY     301 TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
      |||
Db     300 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCTACTGTGAGTACATCAGT 659

QY     361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
      |||
Db     560 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 719

QY     421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
      |||
Db     560 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 779

QY     481 TTTTCTAATCAATTTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
      |||
Db     580 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 839

QY     541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
      |||
Db     640 TCAGATGCCTTCGCTACTTTTAAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 899

QY     601 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
      |||
Db     700 TTCTTAGAACAAAATTATGACACTATTTTGAAGACTATGAGAAACTGCTGCAATCTGAG 959

QY     661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
      |||
Db     760 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 1019

QY     721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAAC 780
      |||
Db     1020 AATTTCAACATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1079

QY     781 CTCCFTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
```


Db 1030 CTGCTTCGAGACAAAAGTCCCAACATCCAATTCTGAAGCCTTCCATGTCTTTAAGGTGTTT 1139
 QY 841 GTGGCCAGTCCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAAA 900
 Db 1140 GTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCCAAA 1199
 QY 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCTGCTGAC 960
 Db 1200 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGTCTGAC 1259
 QY 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 Db 1260 GAGAAGAACTACCTGATTAAACAGATTCTGAGACTTGAAGAAAAGCAGCCCCGTGA 1313

RESULT 3

AK053642

LOCUS AK053642 3039 bp mRNA linear HTC 05-DEC-2002

DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:E130116021 product:M025-LIKE PROTEIN homolog [Homo sapiens], full insert sequence.

ACCESSION AK053642

VERSION AK053642.1 GI:26343600

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 50,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 5 (bases 1 to 3039)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: <http://genome.gsc.riken.go.jp/>


```

FEATURES             Location/Qualifiers
     source            1. .3039
                        /organism="Mus musculus"
                        /mol_type="mRNA"
                        /strain="C57BL/6J"
                        /db_xref="FANTOM_DB:E130116021"
                        /db_xref="taxon:10090"
                        /clone="E130116021"
                        /tissue_type="eyeball"
                        /clone_lib="RIKEN full-length enriched mouse cDNA library"
                        /dev_stage="0 day neonate"
CDS                     287. .1300
                        /note="unnamed protein product; MO25-LIKE PROTEIN homolog
                        [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
                        98.2%ID, 100%length, match=1002)
                        putative"
                        /codon_start=1
                        /protein_id="BAC35457.1"
                        /db_xref="GI:26343601"
                        /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
                        SLQAMKEILCGTNDKEPPTAEVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
                        NILRRQIGTRCPTVEYIISHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
                        FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
                        SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
                        FKVVFVASPHKTQPIVEILLKNQPKLIEFLSSFPQKERTDDEQFADEKNYLIKQIRDLKK
                        AAP"
     polyA_signal      3022. .3027
                        /note="putative"
     polyA_site        3039
                        /note="putative"
BASE COUNT            882 a      645 c      676 g      836 t
ORIGIN

```

Query Match 84.9%; Score 860.4; DB 11; Length 3039;
Best Local Similarity 90.5%; Pred. No. 6.6e-176;
Matches 918; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy	1	ATGAAAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC	60
Db	237	ATGAAAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAAATT	346
Qy	61	CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA	120
Db	347	CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA	406
Qy	121	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	180
Db	407	GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG	466
Qy	181	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTG	240
Db	467	CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG	526
Qy	241	ACACTGATAGCTGACCTGCACCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	300
Db	527	ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	586

VERSION AK076758.1 GI:26345637
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
 Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
 Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
 and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 695-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851

REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1377)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES Location/Qualifiers
source 1. .1377
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4930433N18"
/db_xref="MGI:1894867"
/db_xref="taxon:10090"
/clone="4930433N18"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
CDS 287. .1300
/note="unnamed protein product; MO25-LIKE PROTEIN homolog
[Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
98.2%ID, 100%length, match=1002)
putative"
/codon_start=1
/protein_id="BAC36470.1"
/db_xref="GI:26345638"
/db_xref="MGI:1914081"

/translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
 SLQAMKEILCGTNDKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
 NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECTRHEPLAKIIL
 FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
 SENYVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
 FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLLK
 AAP"

BASE COUNT 430 a 294 c 306 g 347 t
 ORIGIN

Query Match 84.7%; Score 858.8; DB 11; Length 1377;
 Best Local Similarity 90.4%; Pred. No. 1.3e-175;
 Matches 917; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY	1	ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAATCCAGCAGAAATTGTGAAAAATC	60
DB	287	ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAATCCAGCAGAAATTGTCAAAAT	346
QY	61	CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA	120
DB	347	CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA	406
QY	331	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAACAA	180
DB	407	GAGGIGTCAAAATCTCTGCAAGCAATGAAGGAATTCTGTGTGGAACGAACGACAAGGAG	466
QY	332	CCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG	240
DB	467	CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG	526
QY	341	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	300
DB	527	ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	586
QY	351	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
DB	537	TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCTACTGTGAGTACATCAGT	646
QY	361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
DB	547	TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA	706
QY	421	CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC	480
DB	707	CGCTGTGGGATTATGCTAAGAGAGTGTATTGACATGAGCCACTTGCCAAAATCATCCTA	766
QY	431	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
DB	767	TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTGATATCGCT	826
QY	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC	600
DB	827	TCAGATGCCTTCGCTACTTTTAAAGGATTTGTAAACCAGACATAAAGTATTGGTAGCAGAC	886
QY	601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	660
DB	887	TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG	946

QY	661	AATTATGTTACTAAGAGACAGTCTTTAAAGTGCTAGGGGAGCTGATCCTGGACCGTAC	720
Db	947	AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC	1006
QY	721	AACCTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db	1007	AATTTACCAATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC	1066
QY	781	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840
Db	1067	CTGCTTCGAGACAAAAGTCCCAACATCCAATTCTGAAGCCTTCCATGTCTTTAAGGTGTTT	1126
QY	941	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCCAA	900
Db	1127	GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTAAAAAATCAGCCCCAA	1186
QY	901	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGAC	960
Db	1187	CTCATTGAGTTTTTTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC	1246
QY	961	GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA	1014
Db	1247	GAGAAGAACTACCTGATTAAACAGATTTCGAGACTTGAAGAAAAGCAGCCCCGTGA	1300

RESULT 2

BR013205

LOCUS AK013205 1449 bp mRNA linear HTC 05-DEC-2002

DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810430N08 product:MO25-LIKE PROTEIN homolog [Homo sapiens], full insert sequence.

ACCESSION AK013295

VERSION AK013205.1 GI:12850419

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PCBMED 10349636

REFERENCE

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

EURMED 11042159

REFERENCE

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

- Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
- TITLE** RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
- JOURNAL** Genome Res. 10 (11), 1757-1771 (2000)
- MEDLINE** 20530913
- PUBMED** 11076861
- REFERENCE** 4
- AUTHORS** Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bulb, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
- TITLE** Functional annotation of a full-length mouse cDNA collection
- JOURNAL** Nature 409 (6821), 685-690 (2001)
- MEDLINE** 21085660
- PUBMED** 11217851
- REFERENCE** 5
- AUTHORS** The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
- TITLE** Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
- JOURNAL** Nature 420, 563-573 (2002)
- REFERENCE** 6 (bases 1 to 1449)
- AUTHORS** Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
- TITLE** Direct Submission
- JOURNAL** Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT

Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 7.5 and subtraction to
Rot = 37.5. Second strand cDNA was prepared with the primer adapter
of sequence [5'
GAGAGAGAGATCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES

source

Location/Qualifiers

1..1449

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:2810430N08"

/db_xref="MGI:1893512"

/db_xref="taxon:10090"

/clone="2310430N08"

/tissue_type="whole body"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="10, 11 days embryo"

note

281..1292

/note="MO25-LIKE PROTEIN homolog [Homo sapiens]"

(SWISSPROT|Q9H9S4, evidence: FASTY, 98.2%ID, 100%length,
match=1002)

putative"

/db_xref="MGI:1914081"

BASE COUNT

453 a 304 c 325 g 367 t

ORIGIN

Query Match 83.3%; Score 844.8; DB 11; Length 1449;
Best Local Similarity 90.2%; Pred. No. 1.4e-172;
Matches 915; Conservative 0; Mismatches 97; Indels 2; Gaps 1;

QY 1 ATGAAAAAATGCCTTTGTTTGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
|||
Eb 281 ATGAAAAAATGCCCTTGTGTTTGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 340
CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 120
|||
Db 341 CTGAAAGACAACCTGCCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 400
GAAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAAGAA 180
|||
Db 401 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 460
|||
QY 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTG 240

Db	461	CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGCTTGCTGGTG	520
Qy	241	ACACTGATAGCTGACCTGCAGCTGATAGACTTTTGAGGGAAAAAAGATGTGACCCAGATA	300
Db	521	ACACTCATAGCTGACCTGCAGCTCATAGACTTTTGAGGGAAAAAAGAGATGTGACCCAGATA	580
Qy	301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db	581	TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCTACTGTTCGAGTACATCAGT	640
Qy	361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
Db	641	TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA	700
Qy	421	CGTTGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTC	480
Db	701	CGCTGTGGGATTATGCTAAGAGAGTGTATTTCGACATGAGCCACTTGCCAAAATCATCCTA	760
Qy	481	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db	761	TTTTCTAATCAGTTCAGAGATTTCTTCAAGT--GTTGAGCTGTCCACCTTTGATATCGCT	818
Qy	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC	600
Db	819	TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC	878
Qy	601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	660
Db	879	TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAAACTGCTGCAATCTGAG	938
Qy	661	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	720
Db	939	AACTATGTGACAAAGAGACAACTTTAAAGTTTCTAGGTGAGCTGATCCTGGACCGCCAC	993
Qy	721	AACTTTGCCATCATGACAAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db	999	AATTTACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC	1058
Qy	781	CTCCTTCGGGATAAAAAGTCCCAACATCCAAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840
Db	1059	CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT	1118
Qy	841	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA	900
Db	1119	GTGGCCAGCCCCACAAACGCAGCCTATCGTGGAGATTCTGTTAAAAAATCAGCCCCAA	1178
Qy	901	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	960
Db	1179	CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC	1238
Qy	961	GAGAAGAATACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA	1014
Db	1239	GAGAAGAATACTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA	1292

AU125107
 LOCUS AU125107 822 bp mRNA linear EST 01-AUG-2002
 DEFINITION AU125107 NT2RM4 Homo sapiens cDNA clone NT2RM4001047 5', mRNA
 sequence.
 ACCESSION AU125107
 VERSION AU125107.1 GI:10949823
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 822)
 AUTHORS Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J.,
 Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and
 Isogai, T.

TITLE HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S.,
 Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki,
 Y., Sugano, S., Isogai, T.)

JOURNAL Unpublished

COMMENT Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES Location/Qualifiers
 source 1..822
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2RM4001047"
 /cell_type="teratocarcinoma"
 /cell_line="NT2"
 /clone_lib="NT2RM4"
 /note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
 precursor cells"

BASE COUNT 268 a 164 c 171 g 216 t 3 others
 ORIGIN

Query Match 76.0%; Score 770.6; DB 9; Length 822;
 Best Local Similarity 98.5%; Pred. No. 1.5e-156;
 Matches 798; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
 |||
 Db 1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
 QY 79 ATTTTGGAAAA3CAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
 |||
 Db 61 ATTTTGGAAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120
 QY 139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTG 198


```

      |||
Db    121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTG 180
      |||
QY    199 GCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
      |||
Db    181 GCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240
      |||
QY    259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
      |||
Db    241 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300
      |||
QY    319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
      |||
Db    301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360
      |||
QY    379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
      |||
Db    361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420
      |||
QY    439 AGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
      |||
Db    421 AGAGAATGTATTTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480
      |||
QY    499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATAATTGCTTCAGATGCCTTTGCTACT 558
      |||
Db    471 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATAATTGCTTCAGATGCCTTTGCTACT 540
      |||
QY    559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
      |||
Db    541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600
      |||
QY    619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
      |||
Db    601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660
      |||
QY    679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
      |||
Db    661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720
      |||
QY    739 AAGTATATCAGCAAGCCGGAGAACCTGAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
      |||
Db    721 AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 778
      |||
QY    799 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828
      |||
Db    779 CCCAACATCCAGTTTGAAGCCTTCTGGTTTT 808

```

RESULT 7

BX393735

LOCUS BX393735 1201 bp mRNA linear EST 13-MAY-2003

DEFINITION BX393735 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC002YI01 5-PRIME, mRNA sequence.

ACCESSION BX393735

VERSION BX393735.1 GI:30624044

KEYWORDS EST.

SCURCE Homo sapiens (human)

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1201)
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL	Full-length cDNA libraries and normalization
COMMENT	Unpublished
	Contact: Genoscope
	Genoscope - Centre National de Sequencage
	BP 191 91006 EVRY cedex - France
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
	Library was constructed by Life Technologies, a division of
	Invitrogen. This sequence belongs to sequence cluster 6951.r For
	more information about this cluster, see
	http://www.genoscope.cns.fr/
	cgi-bin/cluster.cgi?seq=CS0DC002AE01QP1&cluster=6951.r . Contact :
	Feng Liang Email : fliang@lifetech.com URL :
	http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
	Faraday Avenue Genoscope sequence ID : CS0DC002AE01QP1.

Query Match 74.0%; Score 750.8; DB 13; Length 1201; %
Fast Local Similarity 91.0%; Pred. No. 3.1e-152;
Matches 766; Conservative 45; Mismatches 27; Indels 4; Gaps 2;

QY 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 |||
 Db 623 TTTTACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 682
 QY 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 |||
 Db 583 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 742
 QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
 |||
 Db 743 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 802
 QY 481 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||
 Db 803 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 862
 QY 541 TCAGATGCCTTTTGCTACTTTCAAGGATTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 |||
 Db 663 TCAGATGCCTTTTGCTACTTTCAAGGATTWCTAACCAGACATAAAGTGTGGTAGCAGAC 922
 QY 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 |||
 Db 923 TTCTTAGAACAAAATTACGACACTATTTTTKWWGAYTATGAGAAATTGCTTCAGTCTGAG 982
 QY 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTA-GGGGAGCTGATCCTGGACCGTCA 719
 |||
 Db 983 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTHGGGGGRGTGATCCTGGACCGTCA 1042
 QY 720 CAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAA 779
 |||
 Db 1043 CAACTTTGCCATCATGACAAAGTWTATYMYCYMECCGGSBYHCCYSWWACTCMTGATGAA 1102
 QY 780 CCTCCTTCGGGATAAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTT 839
 |||
 Db 1103 CCCTCYTCGGGGTAAAR---KCCCAMAWCCAGTTTGAAGCCTTTWTKTTTTTWKKGTTTT 1159
 QY 840 TG 841
 |||
 Db 1160 TG 1161

RESULT 8

AK005323

LOCUS

DEFINITION

AK005323 1379 bp mRNA linear HTC 05-DEC-2002
 Mus musculus adult male cerebellum cDNA, RIKEN full-length enriched library, clone:1500031K13 product:MO25-LIKE PROTEIN homolog [Homo sapiens], full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK005323
 AK005323.1 GI:12837793
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1
 Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaïdo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 REFERENCE 6 (bases 1 to 1379)
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,

Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
 Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
 Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
 Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
 Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
 Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
 Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
 Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
 Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
 Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
 Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JCURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
 URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)

COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. Second
 strand cDNA was prepared with the primer adapter of sequence[5'
 GAGAGAGAGAGCGGCCGCAATTAATTCTCGAGTTAATTAAATTAATCCCCCCCCCCC 3']. cDNA
 was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3'
 end: SstI. Host: SOLR.

Location/Qualifiers

1. .1379
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:1500031K13"
 /db_xref="MGI:1901050"
 /db_xref="taxon:10090"
 /clone="1500031K13"
 /sex="male"
 /tissue_type="cerebellum"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 285. .1175
 /note="unnamed protein product; MO25-LIKE PROTEIN homolog
 [Homo sapiens] (SWISSPROT|Q9H9S4, evidence: FASTY,
 98.2%ID, 100%length, match=1002)
 putative"
 /codon_start=1
 /protein_id="BAB23953.2"
 /db_xref="GI:26342524"
 /db_xref="MGI:1916258"
 /translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
 PLQAMKEILCGTNDKEPPTTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN"

NILRRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
SENYVTKRQSLKLLGELILDRHNTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
FKNSVFITNRIHGLKRWLSS"

polyA_signal 1361. .1366
/note="putative"

polyA_site 1379
/note="putative"

BASE COUNT 452 a 284 c 291 g 352 t
ORIGIN

Query Match 70.0%; Score 709.4; DB 11; Length 1379;
Best Local Similarity 88.9%; Pred. No. 3e-143;
Matches 767; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

```
QY      1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
      |||
Db      285 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 344

QY      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 120
      |||
Db      345 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 404

QY      181 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
      |||
Db      405 GAGGTGTCAAAACCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 464

QY      241 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAGAACCTACAGCAGTGGCCTGCTAGTG 240
      |||
Db      455 CCCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGTTGCTGGTG 524

QY      301 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
      |||
Db      525 ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 584

QY      361 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
      |||
Db      585 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCTACTGTGAGTACATCAGT 644

QY      421 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
      |||
Db      645 TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 704

QY      481 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
      |||
Db      705 CGCTGTGGGATTATGCTAAGAGAGTGTATTGACATGAGCCACTTGCCAAAATCATCCTA 764

QY      541 TTTTCTAATCAATTTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
      |||
Db      765 TTTTCTAATCAGTTCAGAGATTTCTTTAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 824

QY      601 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 500
      |||
Db      825 TCAGATGCCTTCGCTACTTTTCAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 884

QY      601 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
      |||
```


[illegible]

RESULT 9

BG218735

LOCUS

BG218735

784 bp mRNA linear EST 21-APR-2001

DEFINITION

EST38476 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.

ACCESSION

3G218735

VERSION

BG218735.1 GI:13744756

KEYWORDS

EST.

FORCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 784)

AUTHORS

1 (bases 1 to 784)
Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
Lerner, L., Costanzo, D., McElligott, K., Boozer, S., Mays, R., Smith,
E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,
J., Danzig, J. and Ducar, M.

TITLE

,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression

JOURNAL

Nat. Biotechnol. 19 (5), 440-445 (2001)

MEDLINE

21227151

PURMED

11329013

CONTENTS

Contact: Scott J. Cain

Athersys, Inc.

3201 Carnegie Ave, Cleveland, OH 44115, USA

Tel: 216 431 9900

Fax: 216 361 9596

Email: scain@athersys.com

High quality sequence stop: 515.

FEATURES

Location/Qualifiers

SOURCE

1. .784

```
/organism="Homo sapiens"
```

```
/mol type="mRNA"
```

```
/db xref="taxon:9606"
```

```
/cell_line="HT1080"
```

```
/clone lib="Athersys RAGE Library"
```

```
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',
```


BASE COUNT	265 a	151 c	158 g	209 t	1 others
ORIGIN					

Query Match 56.3%; Score 671.8; DE 10; Length 784;
Best Local Similarity 96.5%; Pred. No. 3.8e-135;
Matches 718; Conservative 0; Mismatches 23; Indels 3; Gaps 3;

Qy	1	ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC	60
Db	32	ATGAAAAGAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGTAAATC	91
Qy	61	CTGAAAGACAATTGGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA	120
Db	92	CTGAAAGACAATTGGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA	151
Qy	121	GAAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	180
Db	152	GAAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	211
Qy	201	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG	240
Db	202	CCCCCAACAGAAGCAGTGGCCCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG	271
Qy	301	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	360
Db	302	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	331
Qy	381	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db	382	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	391
Qy	401	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCCTTA	420
Db	392	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCCTTA	451
Qy	421	CGTTCTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC	480
Db	432	CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC	511
Qy	481	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db	512	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	571
Qy	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC	600
Db	572	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC	631
Qy	601	TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATG-AGAAATTGCTTCAGTCTGA	659
Db	632	TTCTTAAACAAAATTACGACACTATTTTTGAAGACTATGAAGAAATTGCTTCAGTCTGA	691
Qy	660	GAATTATGTTAC-TAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTC	718

Db 692 GAATTATGTTACTTAAGAGACAGTCTTTAGAGCTGCTAGGGGAGCTGATCCTGAAANGTT 751

Qy 719 ACAACTTTGCCATCATGACAAAGT 742
 |||||

Db 752 ACAACTTTGCC-TCATGACAAAGT 774

RESULT 10

AK013161 1281 bp mRNA linear HTC 05-DEC-2002
 LOCUS
 DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810425013 product:MO25-LIKE PROTEIN homolog [Homo sapiens], full insert sequence.

ACCESSION AK013161
 VERSION AK013161.1 GI:12850350
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE 2

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

REFERENCE 4

AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851

REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1281)

AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Muraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Masukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Moya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' CAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

Location/Qualifiers

1. .1281

```
/organism="Mus musculus"
```

```
/mol type="mRNA"
```

```
/strain="C57BL/6J"
```

```
/db xref="FANTOM_DB:2810425013"
```

```
/db_xref="MGI:1908997"
```

```
/db_xref="taxon:10090"
```

```
/clone="2810425013"
```

```
/tissue type="whole body"
```

```
/tissue_type="whole body  
/clone_lib="RIKEN full-length enriched mouse cDNA library"
```

```
/dev stage="10, 11 days embryo"
```

289. .1096

289. .1096
/note="MO25-LIKE PROTEIN homolog [Homo sapiens]

/note="MO25-LIKE PROTEIN homolog [Homo sapiens]
(SWISSPROT|Q9H9S4, evidence: FASTY, 98.2%ID, 100%length,

```
(SWISEPROF|
match=1002)
```

```
match=100
putative"
```

```
/db xref="MGI:1922871"
```

333 a 275 c 291 g 332 t

305 a 274 5

Query Match 61.4%; Score 622.8; DB 11; Length 1281;

Query Match	81.4%;	Score 0.2270;
Best Local Similarity	82.8%;	Pred. No. 1.7e-124;

Best Local Similarity 82.8%; Pred. NO. 1.7e-124;
Matches 763; Conservative 0; Mismatches 97; Indels 62; Gaps 2;

Qy 93 AGACAAAAAGACAGACAAGGCCTCAGAAGAAGTGTCTAATTCACGCGTTGGTCTT

DB 237 AGAAGACAGGATTTCTAAGGCCTCAGAAGAGGTGTCAAAATCTCTGCAASCHAFISSEH 440

153 AATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGC TCAGCTAGCCCA 212

AATTCGTGTGGAACGAACGACAAGGAGCCCCCTACAGAAGCAGTGGCTCAGCTCCCGA

AC3 AGAACTCTACAGCAGTGGCC TGC TAGTGACAC TGATAGCTGACC TCCTCTCTCT

557 GGAGCTCTACAGCAGCGGGTTGCTGGTGACACTCATAGCTGACCTGCAGCTCATAGACTT 410

QY 273 TGAGGGAAAAAAGATGTGACCCAGATATTTAACCAACAATCTTGAGAAGACACAAATSSS

417 TGAGGGCAAAAAAAGATGTGACCCAGATATTCACCAACATCCCGAGAAGACAGATTCCTTC

Qy. Q33 TCGGAGTCCTACTGTGGAGTATATAAGIGCTCATCCCTCATAACCTCTATTACCTCTTTT

133. ACCGCTGCTCCTACTGTCGAGTACATCAGTTCTCATCCTCACATCCTGTTTATGCTTCTCAA 536

1000 ACCGATATCAACCGCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTTCG 452

|||||
 595 -GGGTATGAAAGCGGCAGAGATTGCCTTACGCTGTGGGATTATGCTAAGAGAGTGTATTTCG 596

512

[illegible]

572

[illegible]

QY 573 AACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGA 632
 Db 709 -----ATTTTTGA 716
 QY 632 AGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCT 692
 Db 717 AGACTATGAGAACTGCTGCAATCTGAGAACTATGTGACAAAGAGACAATCTTTAAAGTT 776
 QY 693 GCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAA 752
 Db 777 GCTAGGTGAGCTGATCCTGGACCGCCACAATTTACCATTATGACCAAG--TATCAGCAA 834
 QY 753 GCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTT 812
 Db 835 GCCAGAGAACCTGAAACTGATGATGAACCTGCTTCGAGACAAAAGTCCCAACATCCAATT 894
 QY 813 TGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGT 872
 Db 895 CGAAGCCTTCCATGTCTTTAAGGTGTTTGTGGCCAGCCCCACAAAACGCAGCCTATCGT 954
 QY 873 GGAGATCCTGTTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGCAGCTTCCAAAAGA 932
 Db 955 GGAGATTCTGTTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGCAGCTTTCAGAAAGA 1014
 QY 933 AAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGA 992
 Db 1015 AAGGACAGACGACGAGCAGTTTGCTGACGAGAAGAACTACCTGATTAAACAGATTCGAGA 1074
 QY 993 CTTGAAGAAAAAGGCCCTTGA 1014
 Db 1075 CTTGAAGAAAGCAGCCCCGTGA 1096

RESULT 11

BU116522

LOCUS BU116522 951 bp mRNA linear EST 25-NOV-2002

DEFINITION 603139786F1 CSEQCHL15 Gallus gallus cDNA clone ChEST129122 5', mRNA sequence.

ACCESSION BU116522

VERSION BU116522.1 GI:25323402

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 951)

AUTHORS Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

TITLE A Comprehensive Collection of Chicken cDNAs

JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)

MEDLINE 22335534

PUBMED 12445392

COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

FEATURES

SOURCE

1. .951

```
/mol type="mRNA"
```

```
/db xref="taxon:9031"
```

```
/sex="Female"
```

```
/tissue type="cerebrum"
```

```
/dev stage="adult"
```

```
/lab host="DH10B"
```

```
/clone lib="CSEQCHL15"
```

```
/note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
```

EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)

[Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally

T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3'

end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,3

end DNA sequencing by GIBCO BRL, Gaithersburg, MD, USA (100000-100000, 624). Cut pBluescript II KS(+) with NotI and EcoRI.

Ligate in double stranded adaptor containing BsgI and

BamHI sites [5'ggccgcgtgcagccccggatccgaaaaaaag]

[5' aattcttttttccggatccggggctgcacgc] "

PAGE FOUR

303 a 206 c 202 g 240 t

ORIGIN

Query Match 60.6%; Score 614; DB 13; Length 951;
Best Local Similarity 32.8%; Pred. No. 1.3e-122;
Matches 737; Conservative 0; Mismatches 150; Indels 3; Gaps 3;

121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 130

6 GAAGTGTCAAATCTCTGCAAGCAATGAAGTAAATTCTGTGTGGGACCACAGACAAGGAG 65

Qy 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
|| || || || || || || || || || || || || || || || || || || || || || || ||

5b 66 CCACCGACAGAAGTAGTGGCTCAGCTGGCACAAGAATTGTACAACAGTGGCCTTCTAGTG 125

241 A C A C T G A T A G C T G A C C T G C A G C T G A T A G A C T T T G A G G G A A A A A A G A T G T G A C C C A G A T A 300

Db 126 A C A C T T A T T G C C A A C C T G C A G C T C A T A G A T T T T G A G G G T A A A A A G G A T G T T C C C A G A T A 185

301. TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCTACTGTGGAGTATATTAGT 360

Db: 186 TTTAACAACATCCTGAGAAGACAAATTGGCACACGAAGCCCTACTGTGGAATACATTAGT 245

QY 561 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420

Db 246 GCCCATCCACATATCCTGTTTCATGCTTCTGAAAGGCTATGAATCCCCAAATATTGCCTTA 305

Qy 421 CGTTGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAA-AATCATCCT 479

db 306 CGCTGTGGAATTATGCTGAGGGAGTGCATCCGACATGAACCATTGGCCAACAATCATACT 365

Qy		180	CTTTTCAATTCAGATTCTTTTAAGTACGTGGAGTTTGTCACACATTTGATATTGC	539
Db		366	TTTTTCAGAACAGTTCAGAGACTTCTTCAAGTATGTGGAAATGTCAACATTTGATATAGC	425
Qy		540	TTCAGATGCCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGA	599
Db		426	ATCTGATGCCCTTTGCTACATTCAAGGACTTGTTAACAAGGCACAAGTTGTTGGTAGCAGA	485
Qy		600	CCTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGA	659
Db		486	TTTTATGGAACAAAATTATGATACGATCTTTGAGGATTATGAAAACTCCTTCATTCTGA	545
Qy		560	GAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCA	719
Db		546	GAATTACGTAACAAAGAGACAGTCTTTGAAGCTGCTGGGTGAATTGATTCTAGACAGACA	605
Qy		720	CAACTTTGCCATCATGACAAAGTATATCAGCAAAGCCGGAGAACCTGAAACTCATGATGAA	779
Db		606	CAACTTCGCCATCATGACAAAATATATCAGCAAACCAGAGAACTCGAAGCTGATGATGAA	565
Qy		780	CCTCCTTCGGGATAAAAAGTCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTT	839
Db		656	CTTGCTGCGAGACAAAAGCCCCAACATTCAATTTGAAGCATTCATGTGTTCAAGGTTT	725
Qy		840	TGTGGCCAGTCCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA	899
Db		726	TGTGGCCAGTCCAAACAAAACCTAGCCCATCGTGGAGATCCTGCTGAAAAACAG-CCAA	784
Qy		900	ACTCATTGAGTTTCTGAGCAGCTTCCA-AAGAGAAAGGACGGATGATGAGCAGTTTCGCYG	958
Db		785	GCTCATCGAGTTTCTGAGCCATTTCAGAAAAGAGAGGACGGTTGACGAGCAGTTCACCG	844
Qy		959	ACGAGAAGAAGTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCC	1008
Db		845	ACGAGAAGAAGTACTTGATCAAGCAAATCCGAGACTTGAAGAAGGCCGAC	994

RESULT 12

10669953

LOCUS B0669953 982 bp mRNA linear EST 15-JUL-2002

DEFINITION AGENCOURT_8203755 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6255924
5' mRNA sequence.

ACCESSION B0669953

VERSION 30669953.1 GI:21780787

KEYWORDS: EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 982)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cqapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2407 row: m column: 13
 High quality sequence stop: 508.

FEATURES

source

Location/Qualifiers

1. .982

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6255924"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_102"

/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed

by Ling Hong in the laboratory of Gerald M. Rubin

(University of California, Berkeley) using ZAP-cDNA

synthesis kit (Stratagene) and Superscript II RT (Life

Technologies). Note: this is a NIH_MGC Library."

BASE COUNT

217 a 200 c 357 g 197 t 11 others

ORIGIN

Query Match 58.6%; Score 594.2; DB 13; Length 982;
 Best Local Similarity 96.5%; Pred. No. 2.6e-118;
 Matches 523; Conservative 0; Mismatches 20; Indels 3; Caps 2;

```

QY 334 CCGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAA 393
    |||
DB 1 CCGAGTCCTACTGTGGAG-ATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAA 59
QY 394 GGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCTGA 453
    |||
DB 50 GGATATGAAGCCCCACAGATTGCCTTACATTGGGGATTATGCTGAGAGAATGGATTCTGA 119
QY 454 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 513
    |||
DB 220 CATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTAC 179
QY 514 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 573
    |||
DB 180 GTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTA 239
QY 574 ACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 633
    |||
DB 240 ACCAGACATAAAGTGGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAA 299
QY 634 GACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTG 693
    |||
DB 300 GACTATGAGAAATTGCTTCAGTCTGAGAATTATGGTACTAAGAGACAGTCTTTAAAGCTG 359
QY 694 CTAGGGGAGCTGATCCTGGACCGTCACAACCTTTGCCATCATGACAAAGTATATCAGCAAG 753
    |||
  
```


Db 360 CTAGGGGAGCTGATCCTGGACCGTCACAACCTTTGCCATCATGACAAAGTATATCAGCAAG 419

QY 754 CCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGTCCCAACATCCAGTTT 813
 |||

Db 420 CCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGTCCCAACATCCAGTTT 479

QY 814 GAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTG 873
 |||

Db 480 GAAGCCTTTCATGTTTTTAAGGGGGTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTG 539

QY 874 GAGATCCTGTAAAAAATCAGCCCCAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAA 933
 |||

Db 540 GAGATCCTGTAAAAAATCAGCCCCAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAA 599

QY 934 AGG--ACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAAC 982
 |||

Db 600 AGGGACGGGATGATGAGCANNTCCCTGACGAGAAAGACTACTTGGGTAAAC 650

RESULT 13

BU518807 934 bp mRNA linear EST 12-SEP-2002

LOCUS BU518807
 DEFINITION AGENCOURT_10171930 NIH_MGC_134 Mus musculus cDNA clone
 IMAGE:6516567 5', mRNA sequence.

ACCESSION BU518807

VERSION BU518807.1 GI:22826333

KEYWORDS EST.

ORIGIN Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 934)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. David Rowe

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM14095 row: e column: 16

High quality sequence stop: 656.

FEATURES

source

Location/Qualifiers

1. .934

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:6516567"

/tissue_type="undifferentiated limb"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_134"

/note="Vector: pCMV-SPORT6.1.cddb; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average

insert size 1.7 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

BASE COUNT 301 a 198 c 200 g 234 t 1 others
ORIGIN

Query Match 57.7%; Score 585.2; DB 13; Length 934;
Best Local Similarity 89.0%; Pred. No. 2.3e-116;
Matches 654; Conservative 0; Mismatches 79; Indels 2; Gaps 2;

```
QY      2  TGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCC 61
      |||
Db      113 TGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATTC 172

QY      62  TGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAG 121
      |||
Db      173 TGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAG 232

QY      122 AAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAAGAAC 181
      |||
Db      233 AGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGCGAACGACAAGGAGC 292

QY      182 CCCCACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTGA 241
      |||
Db      293 CCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTGTGGTGA 352

QY      312 CACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATAT 301
      |||
Db      353 CACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAGATGTGACCCAGATAT 412

QY      363 TTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTG 361
      |||
Db      413 TCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCTACTGTGAGTACATCAGTT 472

QY      423 CTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTAC 421
      |||
Db      473 CTCATCCTCACATCCTGTTTATGCTCCTCAAAGGCTATGAAGCCCCACAGATTGCCTTAC 532

QY      482 GTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTCT 481
      |||
Db      533 GCTGTGGGATTATGCTAAGAGAGTGTATTGACATGAGCCACTTGCCAAAATCATCCTAT 592

QY      542 TTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTT 541
      |||
Db      593 TTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCTT 652

QY      602 CAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACT 601
      |||
Db      653 CAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGACT 712

QY      662 TCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGA 661
      |||
Db      713 TCTTAGAACAAAATTATGACACTATTTTGAAGACTATGAGAAACTGCTGNCATCTGAGA 772

QY      662 ATTATGTTACTAAGAG-ACAGTCTTTAAGCTGCTAGGG-GAGCTGATCCTGGACCGTCA 719
      |||
Db      773 ACTATGTGACAAAGAGAACATTCTTTAAGCTTGCTAGGGTGAGCTGATCCCTGGACCGCC 832
```


Qy 720 CAACTTTGCCATCAT 734
 || || || ||
 Db 833 CACAATTTTCACCAT 847

RESULT 14

CD354831

LOCUS CD354831 713 bp mRNA linear EST 29-MAY-2003

DEFINITION UI-M-GM0-cge-i-10-0-UI.r1 NIH_BMAP_GM0 Mus musculus cDNA clone
 IMAGE: 30361641 5', mRNA sequence.

ACCESSION CD354831

VERSION CD354831.1 GI:31147332

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 713)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

<http://genome.uiowa.edu/distribution/mousefl.html>

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers

1..713

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 30361641"

/tissue_type="whole brain"

/dev_stage="1, 5 and 15 days newborn"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_GM0"

/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;

Site_2: Not I; The library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was size

selected according to mRNA size fraction, ligated with EcoR

I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is CGAACTGAAT. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

BASE COUNT 220 a 159 c 140 g 192 t 2 others
ORIGIN

Query Match 57.1%; Score 579.2; DB 14; Length 713;
Best Local Similarity 89.9%; Pred. No. 4.4e-115;
Matches 642; Conservative 0; Mismatches 70; Indels 2; Gaps 2;

```
Qy 282 AAAAGATGTGACCCAGATATTTAACAACAT-CTTGAGAAGACAGATAGGCACTCGGAGTC 340
    |||
Db 1 AAAAGATGTGACCCAGATATTCAACAACATCCNTGAGAAGACAGATTGGTACACGGTGTG 60

Qy 341 CTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATG 400
    |||
Db 51 CTACTGTGCGAGTACATCAGTTCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATG 120

Qy 401 AAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAAC 460
    |||
Db 121 AAGCCCCACAGATTGCCTTACGCTGTGGGATTATGCTAAGAGAGTGTATTGACATGAGC 180

Qy 461 CACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGT 520
    |||
Db 131 CACTTGCCAAAATCATCCTATTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGC 240

Qy 521 TGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGAC 580
    |||
Db 211 TGTCCACCTTTGATATCGCTTCAGATGCCTTCGCTACTTTTAAGGATTGTGTTAACCAGAC 300

Qy 581 ATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATG 640
    |||
Db 301 ATAAAGTATTGGTAGCAGACTTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATG 360

Qy 641 AGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGG 700
    |||
Db 361 AGAAACTGCTGCAATCTGAGAACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTG 420

Qy 701 AGCTGATCCTGGACCGTCACAACCTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGA 760
    |||
Db 421 AGCTGATCCTGGACCGCCACAATTTCAACATTATGACCAAGTATATCAGCAAGCCAGAGA 480

Qy 761 ACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCT 820
    |||
Db 481 ACCTGAAACTGATGATGAACCTGCTTCGAGACAAAAGTCCCAACATCCAATTGGAAGCCT 540

Qy 821 TTCATGTTTTTAAGGTGTTTGTGGCCAGTCCCTCACAAAACACAGCCTATTGTGGAGATCC 880
    |||
Db 541 TCCATGTCTTTAAGGTGTTTGTGGCCAGCCCCCACAAAACGCAGCCTATCGTGGAGATTG 500

Qy 881 TGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGG 940
    |||
Db 601 TGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAG 660

Qy 941 ATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACT 994
    |||
Db 561 ACGACGAGCAGTTTGCTGACGAGAAG-ACTACCTGATTANACAGATTCGAGACT 713
```


RESULT 15

HSM073180

ID HSM073180 standard; RNA; EST; 742 BP.

XX

AC BX483012;

XX

SV BX483012.1

XX

DT 09-MAY-2003 (Rel. 75, Created)

DT 09-MAY-2003 (Rel. 75, Last updated, Version 1)

XX

DE Homo sapiens mRNA; EST DKFZp686C08234_r1 (from clone DKFZp686C08234)

XX

KW EST; expressed sequence tag.

XX

OS Homo sapiens (human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX

RN [1]

RP 1-742

RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Mewes H.W., Weil B.,

RA Amid C., Osanger A., Fobo G., Han M., Wiemann S.;

RT

RL Submitted (07-MAY-2003) to the EMBL/GenBank/DDBJ databases.

RL MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

XX

CC This is the 5' sequence of the clone insert

CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

CC Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

CC sequenced by MediGenomix (Martinsried/Germany) within the cDNA

CC sequencing consortium of the German Genome Project.

CC No sl sequence available.

CC This clone (DKFZp686C08234) is available at the RZPD in Berlin.

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 5,

CC 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX

EH Key Location/Qualifiers

EH

ET source 1. 742

ET /db_xref="taxon:9606"

ET /mol_type="mRNA"

ET /organism="Homo sapiens"

ET /clone="DKFZp686C08234"

ET /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host

ET DH10B; sites SfiIA + SfiIB"

ET /dev_stage="adult"

ET /tissue_type="cDNA-collection"

XX

SQ Sequence 742 BP; 256 A; 143 C; 162 G; 179 T; 2 other;

Query Match 57.0%; Score 578; DB 2; Length 742;

Best Local Similarity 99.8%; Pred. No. 8e-115;

Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
|||||

Db 164 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 223
 QY 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
 Db 224 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 283
 QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
 Db 284 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 343
 QY 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTG 240
 Db 344 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTG 403
 QY 241 ACACIGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 Db 404 ACACIGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 463
 QY 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 Db 464 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 523
 QY 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 Db 524 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 583
 QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
 Db 584 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 643
 QY 581 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 Db 644 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 703
 QY 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGA 579
 Db 704 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGA 742

Search completed: January 5, 2004, 03:18:09
 Job time : 2589 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 00:37:47 ; Search time 3965 Seconds
(without alignments)
10462.134 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt.....tgaagaaaacggcccttga 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*

28: em_un:*
 29: em_vi:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_pln:*
 35: em_htg_rod:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1014	100.0	1014	6	AX105381	AX105381 Sequence
2	1010.8	99.7	1344	6	AR097361	AR097361 Sequence
3	1010.8	99.7	1344	6	AR203365	AR203365 Sequence
4	1009.2	99.5	1491	9	BC010993	BC010993 Homo sapi
5	992.8	97.9	2002	9	BD157871	BD157871 Primer fo
6	992.8	97.9	2002	9	AK022639	AK022639 Homo sapi
7	860.4	84.9	1359	10	BC016128	BC016128 Mus muscu
8	858.8	84.7	1530	10	BC016546	BC016546 Mus muscu
9	770.6	76.0	822	6	BD147463	BD147463 Primer fo
10	704.6	67.5	831	6	BD079551	BD079551 Cancer-as
11	682.6	57.5	1026	6	AX061831	AX061831 Sequence
12	582.6	57.5	3281	6	AX082322	AX082322 Sequence
13	582.6	57.5	3761	9	BC020570	BC020570 Homo sapi
14	581.6	57.4	1680	9	AF151824	AF151824 Homo sapi
15	581.6	57.4	3466	9	AF113536	AF113536 Homo sapi
16	568.8	56.1	1947	10	BC020041	BC020041 Mus muscu
17	567.2	55.9	2322	10	S51858	S51858 MO25. gene [
18	552.8	54.5	3346	9	AK000804	AK000804 Homo sapi
19	541.6	53.4	1053	6	AX105727	AX105727 Sequence
20	533.6	52.6	851	9	AY211923	AY211923 Homo sapi
21	520.2	51.3	833	6	BD079552	BD079552 Cancer-as
22	478.8	47.2	2366	5	BC044172	BC044172 Danio rer
23	368.8	36.4	2991	10	BC029053	BC029053 Mus muscu
24	365	36.0	158599	2	AC101941	AC101941 Mus muscu
25	362.8	35.8	1894	3	AY128438	AY128438 Drosophil
26	362.8	35.8	8324	2	AC019926	AC019926 Drosophil
27	362.8	35.8	201313	3	AC010688	AC010688 Drosophil
28	362.8	35.8	281101	3	AE003526	AE003526 Drosophil
29	361.2	35.6	2213	3	AB000402	AB000402 Drosophil
30	270	26.6	1452	9	AK026335	AK026335 Homo sapi
31	255	25.1	1466	8	CPR238632	AJ238632 Chlorella
32	240.8	23.7	3514	10	BC034159	BC034159 Mus muscu
33	210.4	20.7	1474	8	AY088359	AY088359 Arabidops

	34	209.6	20.7	1570	8	AF428296	AF428296 Arabidops
	35	200.2	19.7	1032	8	AY055792	AY055792 Arabidops
	36	200.2	19.7	1519	8	AF380659	AF380659 Arabidops
c	37	195	19.2	387	6	AX333971	AX333971 Sequence
c	38	195	19.2	387	6	AX407834	AX407834 Sequence
c	39	186.4	18.4	89137	8	AC006284	AC006284 Arabidops
c	40	185.4	18.3	89807	9	AL138875	AL138875 Human DNA
c	41	185.4	18.3	349980	6	AX711879	AX711879 Sequence
c	42	185.4	18.3	349980	6	AX739961	AX739961 Sequence
c	43	172.8	17.0	159863	9	AL136218	AL136218 Human DNA
	44	169.8	16.7	722	6	AX283846	AX283846 Sequence
	45	166.6	16.4	700	6	AX283847	AX283847 Sequence

ALIGNMENTS

RESULT 1

AX105381 1014 bp DNA linear PAT 30-APR-2001

LOCUS AX105381 Sequence 1 from Patent WO0123552.

ACCESSION AX105381

VERSION AX105381.1 GI:13921508

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS den Daas, I. and Duecker, K.

TITLE Human paralogue of a head trauma induced cytoplasmatic calcium binding protein

JOURNAL Patent: WO 0123552-A 1 05-APR-2001;

MERCK PATENT GmbH (DE)

FEATURES

source

Location/Qualifiers

1..1014

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

CDS

1..1014

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC37735.1"

/db_xref="GI:13921509"

/translation="MKKMPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSK
SLQAMKEILCGTNEKEPPTTEAVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFN
NILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIIL
FSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQ
SENYVTKRQSLKLLGELILDRHNFAMTKYISKPENLKLMMNLLRDKSPNIQFEAFHV
FKVFVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKK
TAP"

BASE COUNT

340 a 205 c 209 g 260 t

ORIGIN

Query Match 100.0%; Score 1014; DB 6; Length 1014;

Best Local Similarity 100.0%; Pred. No. 6.9e-238;

Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
 |||
 Db 1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60

QY 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
 |||
 Db 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120

QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
 |||
 Db 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180

QY 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG 240
 |||
 Db 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG 240

QY 241 ACACCTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 |||
 Db 241 ACACCTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300

QY 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 |||
 Db 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360

QY 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 |||
 Db 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420

QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
 |||
 Db 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480

QY 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||
 Db 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540

QY 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 |||
 Db 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600

QY 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 |||
 Db 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660

QY 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 |||
 Db 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720

QY 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAAC 780
 |||
 Db 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAAC 780

QY 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 |||
 Db 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840

QY 841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 900
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 900
 QY 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTGCTGAC 960
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTGCTGAC 960
 QY 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTGA 1014
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTGA 1014

RESULT 2

AR097361 1344 bp DNA linear PAT 14-FEB-2001
 LOCUS AR097361
 DEFINITION Sequence 2 from patent US 6071721.
 ACCESSION AR097361
 VERSION AR097361.1 GI:12806091
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 UNCLASSIFIED.
 REFERENCE 1 (bases 1 to 1344)
 AUTHORS Tang, Y. Tom., Guegler, K. J., Corley, N. C. and Gorgone, G. A.
 TITLE Calcium binding protein
 JOURNAL Patent: US 6071721-A 2 06-JUN-2000;
 FEATURES Location/Qualifiers
 source 1. 1344
 /organism="unknown"
 BASE COUNT 450 a 261 c 280 g 353 t
 ORIGIN

Query Match 99.7%; Score 1010.8; DB 6; Length 1344;
 Best Local Similarity 99.8%; Pred. No. 4.2e-237;
 Matches 1012; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
 QY 61 CTGAAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 184 CTGAAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243
 QY 121 GAAGTGCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 244 GAAGTGCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303
 QY 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTG 363
 QY 241 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 354 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 423

QY 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 Db 424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483
 QY 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 Db 484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543
 QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
 Db 544 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 603
 QY 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 Db 604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663
 QY 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 Db 664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 723
 QY 601 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 Db 724 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783
 QY 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 Db 784 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843
 QY 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
 Db 844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903
 QY 731 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTAAGGTGTTT 840
 Db 904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTAAGGTGTTT 963
 QY 841 GTGGCCAGTCCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 900
 Db 964 GTGGCCAGTCCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 1023
 QY 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 Db 1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083
 QY 961 GAGAAGAATACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 Db 1084 GAGAAGAATACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1137

RESULT 3

AR203365

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

AR203365

Sequence 2 from patent US 6365371.

AR203365

AR203365.1 GI:21499736

Unknown.

1344 bp

DNA

linear

PAT 20-JUN-2002

ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 1344)
 AUTHORS Tang, Y. Tom., Guegler, K. J., Corley, N. C. and Gorgone, G. A.
 TITLE Calcium binding protein
 JOURNAL Patent: US 6365371-A 2 02-APR-2002;
 FEATURES Location/Qualifiers
 source 1. .1344
 /organism="unknown"
 BASE COUNT 450 a 261 c 280 g 353 t
 ORIGIN

Query Match 99.7%; Score 1010.8; DB 6; Length 1344;
 Best Local Similarity 99.8%; Pred. No. 4.2e-237;
 Matches 1012; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
 Db 124 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
 QY 61 CTGAAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 120
 Db 184 CTGAAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 243
 QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCGTGTGGTACAAACGAGAAAGAA 180
 Db 244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCGTGTGGTACAAACGAGAAAGAA 303
 QY 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
 Db 304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTGGTG 363
 QY 241 AACTGTAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 Db 364 AACTGTAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 423
 QY 301 TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 Db 424 TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483
 QY 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 Db 484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543
 QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
 Db 544 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 603
 QY 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 Db 604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663
 QY 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 Db 664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 723
 QY 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660

Db 724 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

QY 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720

Db 784 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843

QY 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780

Db 844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

QY 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840

Db 904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

QY 941 GTGGCCAGTCTTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCCAA 900

Db 964 GTGGCCAGTCTTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCCAA 1023

QY 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960

Db 1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

QY 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014

Db 1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1137

SEQUENCE

BC010993

LOCUS

DEFINITION

BC010993

1491 bp

mRNA

linear

PRJ 25-JUL-2001

Homo sapiens, hypothetical protein FLJ12577, clone MGC:15031
IMAGE:3956127, mRNA, complete cds.

ACCESSION

BC010993

VERSION

BC010993.1 GI:15012172

KEYWORDS

MGC.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1491)

AUTHORS

Strausberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (23-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbiology.org>

contact: amadan@systemsbiology.org

Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 25 Row: k Column: 12
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10434146.

source

Location/Qualifiers

1. .1491

1. .1491
/organism="Homo sapiens"

```
/mol_type="mRNA"
```

```
/db_xref="LocusID:81617"
```

```
/db_xref="taxon:9606"
```

```
/clone="MGC:15031 IMAGE:3956127"
```

```
/clone="MGC:15031 IMAGE:395012"
/tissue_type="Placenta, choriocarcinoma"
```

```
/tissue_type= Placenta;  
/clone_lib="NIH_MGC_21"
```

```
/clone_11b="N111-M30"
/lab_host="DH10B-R"
```

```
/lab_host="DH10B-K"
/note="Vector: pOTB7"
```

416. .1285

/codon_start=1

```
/codon_start=1
/product="hypothetical protein FLJ12577"

```

```
/product= hypochlorite
/protein_id="AAH10993.1"
```

```
/db_xref="GI:15012173"  
/translation="MKEILCGT"
```

```

/db_xref="GI:15012173"
/translation="MKEILCGTNEKEPPTAEVAQLAQELYSSGLLVTLIADLQLIDFE
EKKDVTQIFNNILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECI

```

EKKDVTQIFNNILRRQIGIRSPVETICAKH
RHEPLAKIILFSNQERDFFKYVELSTEDIASDAFATFKDLLTRHKVLVADFLEQNYDT
LEEDYEKILLOSENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKS
LEEDYEKILLOSENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKS

IFEDYEKLLQSENIVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKDMMNEERL
PNIQFEAFHVFKVVASPHKTPIVEILLKNQPKLIEFLSSFOKERTDDEQFADEKNY

LIKOIRDLKKTAP"

BASE COPY
OBTAIN

503 a 290 c 305 g 393 t

GRANDY MARCH

99.5%; Score 1009.2; DB 9; Length 1491;

```

Query Match          99.5%; Score 1000000
Best Local Similarity 99.7%; Pred. No. 1e-236;
                        0. Mismatches

```

Seq. ID# 1
Query: MARCHA
Best Local Similarity 99.7%; Pred. No. 1e-236;
Matches 1011; Conservative 0; Mismatches 3; Indels 0; Gaps - 0;

Matches 1011; Conservative

Qy 1 ATGAAAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60

|||||

|||||TCTGCAAAATC 33

QY 1 ATGAAAAAAAAATGCCCTTTGGTGTAGTAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 331
| | | | |
| | | | |
DB 279 ATGAAAAAAAAATGCCCTTTGGTGTAGTAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 120

[illegible]

QY 51 CTGAAAGACAAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 391
 DB 322 CTGAAAGACAAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 180

De. 332 CTGAAAGACAAATTTCCG
QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
|||||GCGAAAGAA 451

QY 121 GAAGTGTCTAAATCTAATCTGCAAGCAATGAAAGAAATTCTGTGIGGTACAAACGAGAAAGAA 451
|||
DB 392 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGIGGTACAAACGAGAAAGAA 451
|||||

181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG 240 -

Qy 181 CCCCCAACACAGTCTCAGCTAGCACAAAGACTCTACAGCAGTGGCCTGCTGGTG 511
|||||
Db 452 CCCCAGACAGAAGCAGTGGCTCAGCTAGCACAAAGACTCTACAGCAGTGGCCTGCTGGTG 300

452 CCCCCGACAGAAACCTG
241 A C A C T G A T A G C T G A C C T G C A G C T G A T A G A C T T T G A G G G A A A A A A A G A T G T G A C C C A G A T A 300
QY
573

241 ACAC TGA TAC C T G A C C T G A T A G A C T T T G A G G A A A A A A A G A T G T G A C C C A G A T A 571

512 ACAC TGA TAC C T G A C C T G A T A G A C T T T G A G G A A A A A A A G A T G T G A C C C A G A T A 571

Db 512 ACACCTGATAGCTCGA
Qy 301 TTTAAACAACATCTTGAGAAGACAGATAGGCACTCGGACTCCTACTGTGGAGTATATTAGT 360

Qy

Db 572 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 631
 Qy 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 Db 632 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 691
 Qy 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
 Db 692 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 751
 Qy 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTGATATTGCT 540
 Db 752 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTGATATTGCT 811
 Qy 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
 Db 812 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 871
 Qy 601 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 Db 872 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 931
 Qy 660 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 Db 932 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 991
 Qy 720 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
 Db 992 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 1051
 Qy 780 CTCCTTCGGGATAAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 Db 1052 CTCCTTCGGGATAAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 1111
 Qy 840 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 900
 Db 1112 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 1171
 Qy 900 CTCATTGAGTTTCTGAGCAGCTTCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 Db 1172 CTCATTGAGTTTCTGAGCAGCTTCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1231
 Qy 960 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 Db 1232 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1285

RESULT 5

BD157871

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BD157871

Primer for synthesizing full-length cDNA and use thereof.

BD157871

BD157871.1 GI:27863629

JE 2002191363-A/12714.

Homo sapiens (human)

Homo sapiens

2002 bp

DNA

linear

PAT 17-JAN-2003

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2002)

AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002191363-A 12714 09-JUL-2002;

COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/12714
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SAITO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS (127)..(993).
Location/Qualifiers
FEATURES SOURCE
1..2002
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 594 a 418 c 463 g 527 t
ORIGIN
Query Match 97.9%; Score 992.8; DB 6; Length 2002;
Fast Local Similarity 99.8%; Pred. No. 1.1e-232;
Matches 594; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Q 19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
|
D 1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60
|
Q 79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
|
D 61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120
|
Q 139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 198
|
D 121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 180
|
Q 199 GCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
|
D 181 GCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240
|
Q 259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
|
D 241 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300
|
Q 319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
|

Db 301 AGACAGATAGGCACTCGGAGTCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360

QY 379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438

Db 361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420

QY 439 AGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498

Db 421 AGAGAATGTATTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480

QY 499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558

Db 481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540

QY 559 TTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTAC 618

Db 541 TTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTAC 600

QY 619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678

Db 601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

QY 679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738

Db 661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720

QY 739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798

Db 721 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 780

QY 799 CCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 858

Db 781 CCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 840

QY 859 ACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGC 918

Db 841 ACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGC 900

QY 919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978

Db 901 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960

QY 979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014

Db 961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996

RESULT 6

AK022639

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK022639

Homo sapiens cDNA FLJ12577 fis, clone NT2RM4001047, highly similar to M025 PROTEIN.

AK022639

AK022639.1 GI:10434146

oligo capping; fis (full insert sequence).

Homo sapiens (human)

Homo sapiens

2002 bp

mRNA

linear

PRI 01-AUG-2002

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

1
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,
Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,
Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,
Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A.,
Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
NEDO human cDNA sequencing project

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 2002)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.

FEATURES

SOURCE

Location/Qualifiers

1. 2002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM4001047"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM4"
/note="cloning vector: pME18SFL3-mRNA from uninduced NT2
neuronal precursor cells."

CDS

127..996
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB14147.1"
/db_xref="GI:10434147"
/translation="MKEILCGTNEKEPPTEAVAQLAQELYSSGLLVTLIADLQLIDFE
GKKDVTQIFNNILRRQIGTRSPTVEYISAHPHILFMLLKGYEAPQIALRCGIMLRECI
RHEPLVKIILFSNQFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDT
IFEDYEKLLQSENYVTKRQSLKLLGELILDRHNFAIMTKYISKPENLKLMMNLLRDKS
PNIQFEAFHVFKVVFVSPHKTQPIVEILLKNQPKLIEFLSSFOKERTDDEQFADEKNY
LIKQIRDLKKTAP"

BASE COUNT
ORIGIN

594 a 418 c 463 g 527 t

Query Match 97.9%; Score 992.8; DB 9; Length 2002;
Best Local Similarity 99.8%; Pred. No. 1.1e-232;
Matches 994; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy
Db

19 TTTAGTAAATCACACAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
|||||
1 TTTAGTAAATCACACAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60

Qy 79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
 Db 61 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120

Qy 139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTG 198
 Db 121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTG 180

Qy 199 GCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
 Db 181 GCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240

Qy 259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
 Db 241 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

Qy 319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
 Db 301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360

Qy 379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
 Db 361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420

Qy 439 AGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTGAGA 498
 Db 421 AGAGAATGTATTGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTGAGA 480

Qy 499 GATTTCTTTAAGTACGTTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
 Db 481 GATTTCTTTAAGTACGTTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540

Qy 559 TTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTAC 618
 Db 541 TTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTAC 600

Qy 619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
 Db 601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660

Qy 679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
 Db 661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720

Qy 739 AAGTATATCAGCAAGCCCGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGT 798
 Db 721 AAGTATATCAGCAAGCCCGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGT 780

Qy 799 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 858
 Db 781 CCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 840

Qy 859 ACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGC 918
 Db 841 ACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGC 900

QY 919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
 |||||
 Db 901 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960
 |||||
 QY 979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 |||||
 Db 961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996
 |||||

RESULT 7

BC016128

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK
COMMENT

1359 bp mRNA linear ROD 16-APR-2003
 BC016128
 Mus musculus RIKEN cDNA 1500031K13 gene, mRNA (cDNA clone MGC:28889
 IMAGE:4911640), complete cds.
 BC016128
 BC016128.1 GI:16359341
 MGC.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1359)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, B., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 1359)
 Strausberg, R.
 Direct Submission
 Submitted (22-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 38 Row: m Column: 23
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES

source

Location/Qualifiers

1. .1359
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:28889 IMAGE:4911640"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

gene

1. .1359
/gene="1500031K13Rik"
/note="synonyms: 4930520C08Rik, 2810425013Rik"
/db_xref="LocusID:69008"
/db_xref="MGI:1916258"

CDS

262. .1266
/codon_start=1
/product="1500031K13Rik protein"
/protein_id="AAH16128.1"
/db_xref="GI:16359342"
/db_xref="LocusID:69008"
/translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKKTDKASEEVSKSLQ
AMKEILCGTNDKEPPTTEAVAQLAQELYSSGLLVTLIADLQIDFEGKKDVTQIFNNIL
RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN
YVTKRQSLKLLGELILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
"

BASE COUNT

418 a 301 c 294 g 346 t

ORIGIN

Query Match 84.9%; Score 860.4; DB 10; Length 1359;
Best Local Similarity 90.5%; Pred. No. 3.2e-200;
Matches 918; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

CY

1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
|||||

Db 253 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 312
 QY 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
 Db 313 CTGAAAGACAACCTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 372
 QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
 Db 373 GAGGTGTCAAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG 432
 QY 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
 Db 433 CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG 492
 QY 241 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 Db 493 AACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 552
 QY 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 Db 553 TTCAACAACATCCTGAGAAGACAGATTGGTACACGGTGTCTACTGTGAGTACATCAGT 612
 QY 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 Db 613 GCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA 672
 QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
 Db 673 CGCTGTGGGATTATGCTAAGAGAGTGTATTCGACATGAGCCACTTGCCAAAATCATCCTA 732
 QY 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 Db 733 TTTTCTAATCAGTTCAGAGATTTCTTCAAGTATGTTGAGCTGTCCACCTTTGATATCGCT 792
 QY 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 Db 793 TCAGATGCCTTCGCTACTTTTAAGGATTTGTTAACCAGACATAAAGTATTGGTAGCAGAC 852
 QY 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 Db 853 TTCTTAGAACAAAATTATGACACTATTTTTGAAGACTATGAGAACTGCTGCAATCTGAG 912
 QY 561 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 Db 913 AACTATGTGACAAAGAGACAATCTTTAAAGTTGCTAGGTGAGCTGATCCTGGACCGCCAC 972
 QY 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAAC 780
 Db 973 AATTTACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC 1032
 QY 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 Db 1033 CTGCTTCGAGACAAAAGTCCCAACATCCAATTGCAAGCCTTCCATGTCTTTAAGGTGTTT 1092
 QY 841 GTGGCCAGTCCCTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCCAA 900
 Db 1093 GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTAAAAAATCAGCCCCAA 1152

Qy 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGAC 960
 |||||
 Db 1153 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1212
 Qy 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 |||||
 Db 1213 GAGAAGAACTACCTGATTAAACAGATTCGAGACTTGAAGAAAGCAGCCCCGTGA 1266

RESULT 8

BC016546 1530 bp mRNA linear ROD 16-APR-2003
 LOCUS BC016546
 DEFINITION Mus musculus RIKEN cDNA 1500031K13 gene, mRNA (cDNA clone MGC:27972 IMAGE:3595339), complete cds.
 ACCESSION BC016546
 VERSION BC016546.1 GI:16741456
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1530)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 1530)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (31-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Leulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 35 Row: m Column: 15

This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES

SOURCE

Location/Qualifiers

1. .1530
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:27972 IMAGE:3595339"
/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6".

Gene

1. .1530
/gene="1500031K13Rik"
/note="synonyms: 4930520C08Rik, 2810425013Rik"
/db_xref="LocusID:69008"
/db_xref="MGI:1916258"

CDS

279. .1283
/codon_start=1
/product="1500031K13Rik protein"
/protein_id="AAH16546.1"
/db_xref="GI:16741457"
/db_xref="LocusID:69008"
/translation="MPLFSKSHKNPAEIVKILKDNLAILEKQDKTKASEEVSKSLQ
AMKEILCGTNDKEPPTAEVAQLAQELYSSGLLVTLIADLQLIDFEGKKDVTQIFNNIL
RRQIGTRCPTVEYISSHPHILFMLLKGYEAPQIALRCGIMLRECIRHEPLAKIILFSN
QFRDFFKYVELSTFDIASDAFATFKDLLTRHKVLVADFLEQNYDTIFEDYEKLLQSEN
YVTKRQSLKLRGELIILDRHNFTIMTKYISKPENLKLMMNLLRDKSPNIQFEAFHVFKV
FVASPHKTQPIVEILLKNQPKLIEFLSSFQKERTDDEQFADEKNYLIKQIRDLKKAAP
"

BASE COUNT ORIGIN

498 a 313 c 326 g 393 t

Query Match 84.7%; Score 858.8; DB 10; Length 1530;
Best Local Similarity 90.4%; Pred. No. 8e-200;
Matches 317; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY

1 ATCAAAAAAATGCCTTTGTTTGTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
|||||

Db

270 ATGAAAAAATGCCTTTGTTTGTAGTAAATCACACAAAAATCCAGCAGAAATTGTCAAAATT 329
|||||

QY	51	CTGAAAGACAAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA	120
Db	330	CTGAAAGACAACTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA	389
QY	121	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	180
Db	390	GAGGTGTCAAATCTCTGCAAGCAATGAAGGAAATTCTGTGTGGAACGAACGACAAGGAG	449
QY	181	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG	240
Db	450	CCCCCTACAGAAGCAGTGGCTCAGCTGGCGCAGGAGCTCTACAGCAGCGGGTTGCTGGTG	509
QY	241	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	300
Db	510	ACACTCATAGCTGACCTGCAGCTCATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	569
QY	301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCCTACTGTGGAGTATATTAGT	360
Db	570	TTCAACAACATCTTGAGAAGACAGATTGGTACACGGTGTCTACTGTGCGAGTACATCAGT	629
QY	361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
Db	630	TCTCATCCTCACATCCTGTTTATGCTTCTCAAAGGCTATGAAGCCCCACAGATTGCCTTA	689
QY	421	CGTTGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTC	480
Db	690	CGCTGTGGGATTATGCTAAGAGAGTGTATTTCGACATGAGCCACTTGCCAAAATCATCCTA	749
QY	481	TTTTCTAATCAATTTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db	750	TTTTCTAATCAGTTTCAGAGATTTCTTCAAGTATCTTGAGCTGTCCACCTTTGATATCGCT	809
QY	541	TCAGATGCCTTTGCTACTTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC	600
Db	810	TCAGATGCCTTCGCTACTTTTAAAGGATTTGTAAACCAGACATAAAGTATTGGTAGCAGAC	869
QY	501	TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	560
Db	870	TTCTTAGAACAAAATTATGACACTATTTTGAAGACTATGAGAAACTGCTGCAATCTGAG	929
QY	561	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	620
Db	930	AACTATGTGACAAAAGAGACAATCTTTAAAGTTCGAGGTTGAGCTGATCCTGGACCGCCAC	989
QY	721	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC	780
Db	990	AATTTACCATTATGACCAAGTATATCAGCAAGCCAGAGAACCTGAAACTGATGATGAAC	1049
QY	781	CTCCTTCGGGATAAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840
Db	1050	CTGCTTCGAGACAAAAGTCCCAACATCCAATTCGAAGCCTTCCATGTCTTTAAGGTGTTT	1109
QY	841	GTGGCCAGTCTCTACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAA	900
Db	1110	GTGGCCAGCCCCACAAAACGCAGCCTATCGTGGAGATTCTGTAAAAAATCAGCCCAA	1169

Qy 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGAC 960
 |||
 Db 1170 CTCATTGAGTTTCTGAGCAGCTTTCAGAAAGAAAGGACAGACGACGAGCAGTTTGCTGAC 1229
 Qy 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 |||
 Db 1230 GAGAAGAACTACCTGATTAAACAGATTTCGAGACTTGAAGAAAAGCAGCCCCGTGA 1283

RESULT 9

BD147463

LOCUS BD147463 822 bp DNA linear PAT 17-JAN-2003

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD147463

VERSION BD147463.1 GI:27853221

KEYWORDS JP 2002191363-A/2306.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 822)

AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

TITLE Primer for synthesizing full-length cDNA and use thereof

JOURNAL Patent: JP 2002191363-A 2306 09-JUL-2002;

HEALTH RESEARCH INSTITUTE

COMMENT OS Homo sapiens (human)

BN JP 2002191363-A/2306

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU

PI SAITO,

PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,

PI KEIICHI NAGAI,TETSUJI OTSUKI

PC

C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
 10,

PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key

Location/Qualifiers

FT source 1. .822

FT /organism='Homo sapiens (human)'.
 Location/Qualifiers

FEATURES

source

1. .822

/organism='Homo sapiens'

/mol_type='genomic DNA'

/db_xref='taxon:9606'

BASE COUNT

268 a 164 c 171 g 216 t 3 others

ORIGIN

Query Match 76.0%; Score 770.6; DB 6; Length 822;

Best Local Similarity 98.5%; Pred. No. 3.4e-178;

Matches 798; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

Qy 19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
 |||

Db 1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60

QY 79 ATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
 |||
 Db 61 ATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120
 |||
 QY 139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 198
 |||
 Db 121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTG 180
 |||
 QY 199 GCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
 |||
 Db 181 GCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240
 |||
 QY 259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
 |||
 Db 241 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300
 |||
 QY 319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
 |||
 Db 301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360
 |||
 QY 379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
 |||
 Db 361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420
 |||
 QY 439 AGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGA 498
 |||
 Db 421 AGAGAATGTATTTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTCAGA 480
 |||
 QY 499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
 |||
 Db 481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540
 |||
 QY 559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
 |||
 Db 541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600
 |||
 QY 619 GACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
 |||
 Db 601 GACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660
 |||
 QY 679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
 |||
 Db 661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720
 |||
 QY 739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
 |||
 Db 721 AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 778
 |||
 QY 799 CCCAACATCCAGTTTGAAGCCTTTCATGTT 828
 |||
 Db 779 CCCAACATCCAGTTTGAAGCCTTTCATGTT 808
 |||

RESULT 10
 BD079551
 LOCUS

BD079551

831 bp

DNA

linear

PAT 27-AUG-2002

DEFINITION Cancer-associated nucleic acids and polypeptides.
 ACCESSION BD079551
 VERSION BD079551.1 GI:22625154
 KEYWORDS JP 2001516009-A/217.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 831)
 AUTHORS Old,L.J., Scanlan,M.J., Stockert,E., Gure,A., Chen,Y.T., Gout,I.,
 OGHare,M., Obata,Y., Pfreundschuh,M., Tureci,O. and Sahin,U.
 TITLE Cancer-associated nucleic acids and polypeptides
 JOURNAL Patent: JP 2001516009-A 217 25-SEP-2001;
 LUDWIG INSTITUTE FOR CANCER RESEARCH
 COMMENT OS Homo sapiens (human)
 PN JP 2001516009-A/217
 PD 25-SEP-2001
 PF 15-JUL-1998 JP 2000503425
 PR 17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR
 10-OCT-1997 US 60/061765,10-OCT-1997 US 08/948705 PR
 11-OCT-1997 GB 9721697.2,22-JUN-1998 US 09/102322 PI LLOYD
 J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI TSENG
 CHEN,
 PI IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCHUH, PI
 OZLEM TURECI,
 PI UGUR SAHIN
 EC
 C01N33/574,A61K38/00,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC
 A61P35/00,
 PC C07K14/82,C07K16/32,C12N15/09//C07K16/46,C12P21/08,A61K37/02,
 PC C12N15/00
 CC Cancer-associated nucleic acids and polypeptides..FH Key
 Location/Qualifiers
 FT source 1..831
 FT /organism='Homo sapiens (human)'.
 FEATURES
 source Location/Qualifiers
 1..831
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 285 a 165 c 167 g 209 t 5 others
 ORIGIN

Query Match 67.5%; Score 684.6; DB 6; Length 831;
 Best Local Similarity 96.1%; Pred. No. 4.1e-157;
 Matches 764; Conservative 0; Mismatches 23; Indels 8; Gaps 6;

QY 1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
 |||
 Db 37 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 96
 QY 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 120
 |||
 Db 97 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 156
 QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
 |||

Db 157 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATCTGTGTGGTACAAACGAGAAAGAA 216

Qy 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
 |||

Db 217 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 276
 |||

Qy 241 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 |||

Db 277 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 336
 |||

Qy 301 TTTAACAACATCTTGAGAAGACAGATAGGCACCTCGGAGTCCTACTGTGGAGTATATTAGT 360
 |||

Db 337 TTTAACAACATCTTGAGAAGACAGATAGGCACCTCGGAGTCCTACTGTGGAGTATATTAGT 396
 |||

Qy 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 |||

Db 397 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 456
 |||

Qy 421 CGTTGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACCTTGCCAAAATCATCCTC 480
 |||

Db 457 CGTTGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACCTTGCCAAAATCATCCTC 516
 |||

Qy 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||

Db 517 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 576
 |||

Qy 541 TCAGATGCCTTTGCTACTTTCAA-GGATTTACTAACCAGACATAAAGTGTGGTAGC-AG 598
 |||

Db 577 TCAGATGCCTTTGCTACTTTCAAGGGATTTACTAACCAGACATAAAGTGTGGTAGCAAG 636
 |||

Qy 599 ACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 658
 |||

Db 637 ACTTCTTAGAACAAAATTACGACACTANTTTTGAAGACTATGAGAAATTGCTTCAGTCTG 696
 |||

Qy 659 AG-AATTATGTTACTAAGAGACAGTCTTTAAAG-CTGCTAGGGGAGCTGATCCTGGACCG 716
 |||

Db 697 AGAAATTATGTTACCAAGAGACAGTCTTTAAAGCCTGCTAAGGGAAGTATTCTGGACCG 756
 |||

Qy 717 TCACAACCTTTGCCATC-ATGACAAAGTATATCAGCAAGCC---GGAGAACCTGAAACTCA 772
 |||

Db 757 TCANAACCTTTGCCATCAANGCAAAAGTTTATCAACAAGCCNGGGGAAACCGGAAACNCAA 816
 |||

Qy 773 TGATGAACCTCCTTC 787
 |||

Db 817 GGAGGAACCTCCTTC 831

RESULT 11

AX061831

LOCUS AX061831 1026 bp DNA linear PAT 24-JAN-2001

DEFINITION Sequence 1 from Patent WO0078947.

ACCESSION AX061831

VERSION AX061831.1 GI:12539911

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY 429 GATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
 || ||| | ||||| || ||||| ||||| ||||| || |||
 Db 432 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCCAAAATCATTGTGGTCCGA 491
 QY 439 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
 || || ||||| | ||| || ||||| ||||| ||||| |||||
 Db 492 ACAGTTTTATGATTTCTTCAGATATGTGCAATGTCAACATTTGACATAGCTTCAGATGC 551
 QY 549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA 608
 ||||| || ||||| ||||| || ||||| || ||||| |||||
 Db 552 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTGGGA 611
 QY 569 ACAAATACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
 ||| |||| || | |||| || ||||| || ||||| |||||
 Db 612 ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTACTTCATTACAGAAAATTATGT 671
 QY 669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTTCG 728
 || || ||||| || ||||| || ||||| || ||||| |||||
 Db 672 GACAAAAGACAGTCACTGAAGCTTCTCGGTGAAGTACTACTAGATAGACACAACCTTTCAC 731
 QY 729 CATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
 || ||||| || ||||| || ||||| || ||||| |||||
 Db 732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 791
 QY 789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
 || ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 792 AGACAAAAGTCGCAACATCCAGTTTGAAGCCTTTTCAGTTTTTAAGGTGTTTGTAGCCAA 851
 QY 909 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACCTCATTGA 908
 |||| |||| || ||||| || ||||| || ||||| |||||
 Db 912 TCCTAACAGACGCAGCCCATCTAGACATCCTCTCAAGAACCAGGCCAAACCTCATAGA 911
 QY 909 GTTTCGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
 |||| || |||| || |||| || ||||| ||||| |||||
 Db 912 GTTCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAAACGACGAGAAGAC 971
 QY 969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
 |||| || ||||| || ||||| || |||||
 Db 972 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1012

RESULT 12

AX082322

LOCUS

AX082322

3281 bp

DNA

linear

PAT 28-FEB-2001

DEFINITION

Sequence 26 from Patent WO0111032.

ACCESSION

AX082322

VERSION

AX082322.1 GI:13184499

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Hodgson, D.M., Lincoln, S.E., Russo, F.D., Spiro, P.A., Banville, S.C.,
 Bratcher, S.R., DuFour, G.E., Cohen, H.J., Rosen, B.H., Chalup, M.S.,
 Hillman, J.L., Jones, A.L., Yu, J.Y., Greenawalt, L.B., Panzer, S.R.,

Roseberry, A.M., Wright, R.J. and Daniels, S.E.

TITLE Secretory molecules
JOURNAL Patent: WO 0111032-A 26 15-FEB-2001;
Incyte Genomics, Inc. (US)

FEATURES Location/Qualifiers
source 1. .3281
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 481257.3"

BASE COUNT 1014 a 601 c 676 g 990 t
ORIGIN

Query Match 57.5%; Score 582.6; DB 6; Length 3281;
Best Local Similarity 74.7%; Pred. No. 4.3e-132;
Matches 748; Conservative 0; Mismatches 244; Indels 9; Gaps 1;

```
Qy 18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGC 77
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAAATCTGAAGGAGAGCATGGC 160

Qy 78 CATTTTGGAAAAGCAAGAC-----AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 220

Qy 129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGACCCCAAC 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAATGAAAAAGAGCCTCAGAC 280

Qy 189 AGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 281 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCTGGT 340

Qy 349 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAA 308
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 341 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 400

Qy 309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 401 TATTCTCAGAAGACAAATTTGGTACGAGAATCCTACTGTTGAATACATCTGCACCCAACA 460

Qy 369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 461 GAATATTTTGTTCATGTTATTGAAAGGCTATGAATCTCCAGAAATAGCTCTAAATTGTGG 520

Qy 429 GATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 521 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 580

Qy 489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 581 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 640

Qy 549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGA 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 641 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTCAGAATTTTGGG 700
```


Qy 609 ACAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 568
 Db 701 ACAGCATTATGATAGATTTTTTCAGTGAATATGAGAAGTTACTTCATTTCAGAAAATTATGT 760

Qy 669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
 Db 761 GACAAAAGACAGTCACTGAAGCTTCTCGGTGAAGTACTACTAGATAGACACAACCTTCAC 820

Qy 729 CATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
 Db 821 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 880

Qy 789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
 Db 881 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 940

Qy 849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACTCATTGA 908
 Db 941 TCCTAACAAGACGCAGCCCATCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGA 1000

Qy 909 GTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
 Db 1001 GTTCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1060

Qy 969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
 Db 1061 CTATTAGTTAAACAGATCAGGGATTGAAGAGACCAGCTC 1101

FIGURE 13

BC020570

DEFINITION

3761 bp mRNA linear PRI 22-JAN-2002
 Homo sapiens, MO25 protein, clone MGC:21631 IMAGE:4397573, mRNA, complete cds.

ACCESSION

BC020570

VERSION

BC020570.1 GI:18088260

KEYWORDS

MGC.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 3761)

AUTHORS

Strausberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (03-JAN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Db 576 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 635
 QY 309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC 368
 Db 636 TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA 695
 QY 369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
 Db 696 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 755
 QY 429 GATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
 Db 756 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCCAAAATCATTTTGTGGTCGGA 815
 QY 489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
 Db 816 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 875
 QY 549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGA 608
 Db 876 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAAATTTTGA 935
 QY 609 ACAAATTCAGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
 Db 936 ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTACAGAAATTTATGT 995
 QY 937 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTTCG 728
 Db 996 GACAAAAGACAGTCACTGAAGCTTCTCGGTGAAGTACTACTAGATAGACACAACCTTCAC 1055
 QY 729 CATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
 Db 1056 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1115
 QY 789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
 Db 1116 AGACAAAAGTCGCAACATCCAGTTTGAAGCCTTTTCAGTTTTTAAGGTGTTTGTAGCCAA 1175
 QY 849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGA 908
 Db 1176 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCCAAACCTCATAGA 1235
 QY 909 GTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAA 968
 Db 1236 GTTCCTCAGCAAGTTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1295
 QY 969 CTAATTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
 Db 1296 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1336

RESULT 14

AF151824

LOCUS

AF151824

1680 bp

mRNA

linear

PRI 18-MAY-2000

DEFINITION

Homo sapiens CGI-66 protein mRNA, complete cds.

ACCESSION

AF151824

QY 250 GCTGACCTGCGAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAAC 309
Db 253 GCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAGACGTGGCTCAAATTTTCAACAAT 312
QY 310 ATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCT 369
Db 313 ATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACAG 372
QY 370 CATATCCTGTTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGG 429
Db 372 AATATTTTGTTCATGTTATTGAAAGGTATGAATCTCCAGAAATAGCTCTAAATTGTGGA 432
QY 430 ATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
Db 433 ATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCCGAA 492
QY 490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
Db 493 CAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGCA 552
QY 550 TTTGCTACTTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAA 609
Db 553 TTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTGGAA 612
QY 560 CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT 669
Db 563 CAGGATTAATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTGAGAAAATTATGTG 672
QY 720 ACTAAGAGACAGTCTTTAAAGCTGCTACGGGAGCTGATCCTGGACCGTCACAACCTTTGCC 729
Db 723 ACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACCTACTACTAGATAGACACAACCTCACA 732
QY 730 ATCAIGACAAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGG 789
Db 733 ATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGA 792
QY 790 GATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGT 849
Db 793 GACAAAAGTCGCAACATCCAGTTTGAAGCCTTTCAGGTTTTTAAGGTGTTTGTAGCCAAT 852
QY 850 CCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACCTCATTGAG 909
Db 853 CCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACCTCATAGAG 912
QY 910 TTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAAC 969
Db 913 TTCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAAACGACGAGAAGACC 972
QY 970 TACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
Db 973 TATTTAGTTAAACAGATCAGGGATTGAAAGAGACCAGCTC 1012

AF113536

mRNA

linear

PRI 04-DEC-1999

QY 190 GAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTGACACTGATA 249
 ||||| ||||| || || ||||| ||||| || ||||| |||||
 Db 246 GAAGCAGTAGCTCAACTTGCTCAAGAAGCTCTATAATAGTGGGCTCCTTAGCACCTGGTA 305
 QY 250 GCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAAC 309
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 306 GCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAGACGTGGCTCAAATTTCAACAAT 365
 QY 310 ATCTTGAGAAGACAGATAGGCACTCGGAGTCTACTGTGGAGTATATTAGTGCTCATCCT 369
 || || ||||| || || || || || ||||| || || || || || || ||
 Db 366 ATCTCAGAAGACAAATTGGTACGAGAAGCTCTACTGTTGAATACATCTGCACCCAACAG 425
 QY 370 CATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGG 429
 |||| |||| |||| || ||||| ||||| ||||| ||||| |||||
 Db 425 AATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGGA 485
 QY 430 ATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAAT 489
 || || || ||||| || || ||||| ||||| ||||| ||||| |||||
 Db 486 ATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGAA 545
 QY 490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
 || || ||||| || || || || ||||| ||||| ||||| ||||| |||||
 Db 546 CAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGCA 605
 QY 550 TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAA 609
 |||| || ||||| ||||| || ||||| ||||| ||||| ||||| |||||
 Db 556 TTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTGGAA 565
 QY 570 CAAAATACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT 569
 || ||||| || || || || ||||| ||||| ||||| ||||| |||||
 Db 576 CAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTGAGAAAATTATGTG 725
 QY 580 ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCC 729
 || ||||| ||||| ||||| || ||||| ||||| ||||| ||||| |||||
 Db 735 ACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACTTCACA 785
 QY 730 ATCATGACAAAAGTATATCAGCAAGCCGGAGAAGCTGAAACTCATGATGAACCTCCTTCGG 789
 || ||||| || ||||| || || ||||| ||||| ||||| ||||| |||||
 Db 735 ATTATGACAAAATACATCAGTAAACCTGAGAAGCTCAAATTAATGATGAACCTGCTGCGA 845
 QY 790 GATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGT 849
 || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 845 GACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAAT 905
 QY 850 CCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACCTCATTGAG 909
 || ||||| || ||||| || || ||||| ||||| ||||| ||||| |||||
 Db 906 CCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACCTCATAGAG 965
 QY 910 TTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGACGAGAAGAAC 969
 || || ||||| || || ||||| ||||| ||||| ||||| ||||| |||||
 Db 955 TTCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACC 1025
 QY 970 TACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
 || || ||||| || || ||||| || || || || || || ||
 Db 1026 TATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1065

Search completed: January 6, 2004, 02:34:57
Job time : 3971 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 00:31:27 ; Search time 390 Seconds
(without alignments)
7018.539 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	%		Query		DB	ID	Description
	Score	Match	Length				
1	1014	100.0	1014	22	AAF86462		Human Acute Neuron
2	1014	100.0	1421	22	AAI58234		Human polynucleoti
3	1010.8	99.7	1344	21	AAA27332		Human calcium bind
4	992.8	97.9	2002	22	AAH15879		Human cDNA sequenc
5	770.6	76.0	822	22	AAH05471		Human cDNA clone (
6	684.6	67.5	831	20	AAX39817		Gastric cancer ass
c 7	634.4	67.5	1191	22	AAI60020		Human polynucleoti
8	582.6	57.5	1026	22	AAC91772		Human ANIC-BP (acu
9	582.6	57.5	3281	24	ABK13127		Human secretory po
10	582.6	57.5	3849	23	ABV22987		Human prostate exp
11	582.6	57.5	3849	23	ABV28822		Human prostate exp
12	541.6	53.4	1053	22	AAF30688		Human acute neuron
13	539.6	53.2	1162	23	AAS89557		DNA encoding novel
c 14	520.2	51.3	833	20	AAX39818		Gastric cancer ass
15	496	48.9	2492	23	AAS88031		DNA encoding novel
16	387.8	38.2	722	20	AAZ15133		Human gene express
17	362.3	35.8	2231	23	ABL07151		Drosophila melanog
18	362.3	35.8	4231	23	ABL07150		Drosophila melanog
19	283.8	28.5	690	24	ABS77084		Frog embryonic gen
20	246.4	24.3	435	24	ABL82285		Human ovarian canc
21	244.8	24.1	447	24	ABL82921		Human ovarian canc
22	244.8	24.1	450	24	ABL81975		Human ovarian canc
23	210.3	20.8	762	24	ABS76784		Frog embryonic gen
24	210.4	20.7	1474	21	AAC32983		Arabidopsis thalia
25	208.8	20.6	1497	21	AAC40181		Arabidopsis thalia
26	200.2	19.7	918	21	AAC42766		Arabidopsis thalia
27	200.2	19.7	1032	21	AAC48253		Arabidopsis thalia
c 28	195	19.2	387	24	ABN93983		Gene #481 used to
c 29	195	19.2	387	24	ABL66143		Lung cancer relate
30	169.8	16.7	722	24	AAS61992		Porcine muscular s
31	166.6	16.4	700	24	AAS61993		Porcine muscular s
32	163.8	16.2	481	25	ABZ19574		Group III cDNA can
33	163.4	16.1	300	20	AAZ14552		Human gene express
34	161.2	15.9	1515	21	AAC50415		Arabidopsis thalia
35	156	15.4	861	24	ABN98824		Arabidopsis thalia
c 36	153.4	15.1	737	23	AAS79449		DNA encoding novel
37	147.2	14.5	464	21	AAC46721		Zea mays DNA fragm
38	133.2	13.1	615	22	AAH07116		Human cDNA clone (
39	107.6	10.6	1149	23	AAS88030		DNA encoding novel
40	107.6	10.6	3279	23	AAS89559		DNA encoding novel
41	65.6	6.5	432	24	ABN78107		Human ORF3054 cDNA
42	65	6.4	487	22	AAI98879		Human excretory re
43	65	6.4	487	22	AAI64066		Human bladder rela
44	53.6	5.3	254	25	ABX31310		Human GDP-mannose
45	43	4.2	447	21	AAC06449		Human secreted pro

ALIGNMENTS

RESULT 1

AAF86462

ID AAF86462 standard; cDNA; 1014 BP.

XX

AC AAF86462;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human Acute Neuronal Induced Calcium Binding Protein, ANIC-BP, cDNA.

XX

KW Human; cerebroprotective; neuroprotective; vulnerary; vaccine;
KW gene therapy; Acute Neuronal Induced Calcium Binding Protein; ANIC-BP;
KW stroke; acute head trauma; multiple sclerosis; spinal cord injury; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..1014

FT /*tag= a

FT /product= "Human Acute Neuronal Induced Calcium Binding
FT Protein, ANIC-BP"

XX

PN NC200123552-A1.

XX

DE 05-APR-2001.

XX

DE 13-SEP-2000; 2000WO-EP09132.

XX

DE 04-SEP-1999; 99EP-0118848.

XX

CA (PCT) MERCK PATENT GMBH.

XX

PI Ben Claus I, Duecker K;

XX

DR WPI; 2001-308142/32.

DR

P-PSDB; AAB82090.

XX

PT Novel human acute neuronal induced calcium binding polypeptide, and
PT polynucleotides encoding them useful for diagnosing or treating stroke,
PT acute head trauma, multiple sclerosis and spinal cord injury -

XX

PS Claim 5; Page 40-41; 45pp; English.

XX

CC The present sequence is the coding sequence for human Acute Neuronal
CC Induced Calcium Binding Protein (ANIC-BP). ANIC-BP coding sequence and
CC protein are useful for treating stroke, acute head trauma, multiple
CC sclerosis and spinal cord injury. ANIC-BP coding sequence and protein
CC are also useful as vaccines for inducing an immunological response in a
CC mammal.

XX

SQ Sequence 1014 BP; 340 A; 205 C; 209 G; 260 T; 0 other;

Query Match 100.0%; Score 1014; DB 22; Length 1014;

Best Local Similarity 100.0%; Pred. No. 3.5e-272;

Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
 |||
 Db 1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60

Qy 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
 |||
 Db 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120

Qy 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
 |||
 Db 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180

Qy 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTG 240
 |||
 Db 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTG 240

Qy 241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 |||
 Db 241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300

Qy 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 |||
 Db 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360

Qy 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 |||
 Db 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420

Qy 421 CGTTGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTC 480
 |||
 Db 421 CGTTGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTC 480

Qy 481 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||
 Db 481 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540

Qy 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 |||
 Db 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600

Qy 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 |||
 Db 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660

Qy 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 |||
 Db 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720

Qy 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAAC 780
 |||
 Db 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAAC 780

Qy 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 |||
 Db 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840

Qy 841 GTGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAA 900

Db 841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 900
 QY 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 Db 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 QY 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 Db 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014

RESULT 2

AAI58234

ID AAI58234 standard; cDNA; 1421 BP.

XX

AC AAI58234;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 437.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX

OS Homo sapiens.

XX

PN F0200153312-A1.

XX

PR 26-JUL-2001.

XX

PR 26-DEC-2000; 2000WO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR

PR 25-APR-2000; 2000US-0552317.

PR

PR 09-JUL-2000; 2000US-0598042.

PR

PR 19-JUL-2000; 2000US-0620312.

PR

PR 03-AUG-2000; 2000US-0653450.

PR

PR 14-SEP-2000; 2000US-0662191.

PR

PR 19-OCT-2000; 2000US-0693036.

PR

PR 29-NOV-2000; 2000US-0727344.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-442253/47.

DR

DR P-PSDB; AAM39078.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT

PT such as central nervous system injuries

XX

PS Claim 1; SEQ ID NO 437; 10078pp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX

SQ Sequence 1421 BP; 469 A; 284 C; 306 G; 362 T; 0 other;

Query Match 100.0%; Score 1014; DB 22; Length 1421;
Best Local Similarity 100.0%; Pred. No. 4e-272;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy: 1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 50
|
Db 217 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 276
Qy 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
|
Db 277 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 336
Qy 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
|
Db 337 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 396
Qy 381 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG 240
|
Db 397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG 456
Qy 241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
|
Db 457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 516
Qy 301 TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
|
Db 517 TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
Qy 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
|
Db 577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
Qy 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
|
Db 637 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 696

QY 481 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 697 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 756
 QY 541 TCAGATGCCTTTGCTACTTTCAAGGATTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 757 TCAGATGCCTTTGCTACTTTCAAGGATTACTAACCAGACATAAAGTGTGGTAGCAGAC 816
 QY 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 817 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 876
 QY 561 AATTATGTTACTAAGAGACAGTCTTTAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 877 AATTATGTTACTAAGAGACAGTCTTTAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 936
 QY 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 937 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 996
 QY 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 997 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 1056
 QY 841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 900
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 957 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 1116
 QY 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1117 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1176
 QY 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1177 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1230

RESULT 3

AAA27332

ID AAA27332 standard; cDNA; 1344 BP.

XX

AC AAA27332;

XX

DT 10-AUG-2000 (first entry)

XX

DE Human calcium binding protein hCBP gene.

XX

KW Human; calcium binding protein; cancer; inflammation; CBP;

KW reproductive disorder; autoimmune disorder; developmental disorder;

KW seizure disorder; immune disorder; infection; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 124..1134

FT /*tag= a

FT /product= "calcium binding protein"

XX

PN WO200029580-A1.

XX

PD 25-MAY-2000.

XX

PF 12-NOV-1999; 99WO-US27027.

XX

PR 13-NOV-1998; 98US-0190965.

XX

PA (INCY-) INCYTE PHARM INC.

XX

PI Tang YT, Guegler KJ, Corley NC, Gorgone GA;

XX

DR WPI; 2000-387793/33.

DR P-PSDB; AAY94247.

XX

PT Human hCBP protein, and the nucleic acid encoding it, useful for e.g.

PT diagnosis, prevention and treatment of cancers, immune, developmental

PT or reproductive disorders -

XX

PS Claim 9; Fig 1; 72pp; English.

XX

CC The present sequence is the human calcium binding protein hCBP gene. It
CC was obtained by screening a coronary artery smooth muscle cDNA library,
CC from which five overlapping nucleic acids were isolated and
CC sequenced, and then expressed to give the protein. The protein and the
CC gene encoding it are useful for the diagnosis and treatment of the
CC following types of disorder: cancers (such as adenocarcinomas),
CC reproductive disorders (such as infertility, ovulatory defects),
CC endometriosis, disruptions of the oestrus and menstrual cycles,
CC polycystic ovary syndrome and ovarian hyperstimulation), autoimmune
CC disorders (such as benign prostatic hyperplasia and prostatitis),
CC developmental disorders (such as Cushing's syndrome, muscular dystrophy
CC and gonadal dysgenesis), hereditary neuropathies, seizure disorders,
CC immune disorders (such as AIDS, allergies, anaemia, asthma,
CC atherosclerosis, cholecystitis, Crohn's disease, diabetes, Graves'
CC disease, multiple sclerosis, psoriasis, rheumatoid arthritis,
CC scleroderma, Sjogren's syndrome and ulcerative colitis), and viral,
CC bacterial, fungal, parasitic, protozoal and helminthic infections.

XX

SQ Sequence 1344 BP; 450 A; 261 C; 280 G; 353 T; 0 other;

Query Match 99.7%; Score 1010.8; DB 21; Length 1344;

Best Local Similarity 99.8%; Pred. No. 3.1e-271;

Matches 1012; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
|
Db 124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183
|
QY 61 CTGAAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 120
|
Db 184 CTGAAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 243
|
QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
|

Db	244	GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	303
Qy	181	CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTG	240
Db	304	CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTGGTG	363
Qy	241	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	300
Db	364	ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA	423
Qy	301	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	360
Db	424	TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT	483
Qy	361	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	420
Db	484	GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA	543
Qy	421	CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC	480
Db	544	CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC	603
Qy	481	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	540
Db	604	TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT	663
Qy	541	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC	600
Db	664	TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC	723
Qy	601	TTCTTAGAACAATAATACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	660
Db	724	TTCTTAGAACAATAATACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG	783
Qy	661	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	720
Db	784	AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC	843
Qy	721	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACC'TGAAACTCATGATGAAC	780
Db	844	AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACC'TGAAACTCATGATGAAC	903
Qy	781	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	840
Db	904	CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT	963
Qy	841	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA	900
Db	964	GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA	1023
Qy	901	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	960
Db	1024	CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC	1083
Qy	961	GAGAAGAATACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA	1014
Db	1084	GAGAAGAATACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA	1137

RESULT 4

AAH15879

ID AAH15879 standard; cDNA; 2002 BP.

XX

AC AAH15879;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human cDNA sequence SEQ ID NO:14407.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2003.

XX

PF 28-JUL-2000; 2000EP-0116126.

XX

PR 09-JUL-1999; 99JP-0248036.

PR 07-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 14-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX

REVI-1) RELEX RES INST.

XX

Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto S; Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

WIPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

XX

PS Claim 8; SEQ ID 14407; 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX

SQ Sequence 2002 BP; 594 A; 418 C; 463 G; 527 T; 0 other;

Query Match 97.9%; Score 992.8; DB 22; Length 2002;
 Best Local Similarity 99.8%; Pred. No. 3.8e-266;
 Matches 994; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY      19 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
      |||
Db      1 TTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60

QY      79 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 138
      |||
Db      51 ATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTG 120

QY     139 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTG 198
      |||
Db     121 CAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTG 180

QY     199 GCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
      |||
Db     151 GCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 240

QY     259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
      |||
Db     243 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

QY     319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 373
      |||
Db     301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360

QY     379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
      |||
Db     351 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420

QY     439 AGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTGAGA 498
      |||
Db     421 AGAGAATGTATTTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTGAGA 480

QY     499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
      |||
Db     481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540

QY     559 TTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTAC 618
      |||
Db     541 TTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTAC 600

QY     619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
      |||
Db     601 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660
  
```


QY 679 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720
 QY 739 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGT 798
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 721 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGT 780
 QY 799 CCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 858
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 781 CCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAA 840
 QY 959 ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGC 918
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 941 ACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGC 900
 QY 919 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 978
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 901 AGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATT 960
 QY 979 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 ||||||||||||||||||||||||||||||||||||||||
 Db 961 AAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 996

RESULT 5

AAH05471

ID AAH05471 standard; cDNA; 322 BP.

XX

AC AAH05471;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human cDNA clone (5'-primer) SEQ ID NO:2306.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PE 23-JUL-2000; 2000EP-0116126.

XX

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

XX

PS Claim 1; SEQ ID 2306; 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX

SO Sequence 822 BP; 268 A; 164 C; 171 G; 216 T; 3 other;

Query Match 76.0%; Score 770.6; DB 22; Length 822;
Best Local Similarity 98.5%; Pred. No. 2.1e-204;
Matches 798; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 19 TTTAGTAAATCACACAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 78
|
Db 1 TTTAGTAAATCACACAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCC 60

QY 79 ATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGCTAAATCACTG 138
|
Db 61 ATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGCTAAATCACTG 120

QY 139 CAAGCAATGAAAGAAATCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTG 198
|
Db 121 CAAGCAATGAAAGAAATCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTG 180

QY 199 GCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTG 258
|
Db 181 GCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTGGTGACACTGATAGCTGACCTG 240

QY 259 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 318
|
Db 241 CAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGA 300

QY 319 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 378
 |||
 Db 301 AGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTG 360
 QY 379 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 438
 |||
 Db 361 TTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTG 420
 QY 439 AGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTGAGA 498
 |||
 Db 421 AGAGAATGTATTTCGACATGAACCACTTGTCAAAATCATCCTCTTTTCTAATCAATTGAGA 480
 QY 499 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 558
 |||
 Db 481 GATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT 540
 QY 559 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
 |||
 Db 541 TTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 600
 QY 619 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 678
 |||
 Db 501 GACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGA 660
 QY 739 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 738
 |||
 Db 661 CAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACA 720
 QY 799 AAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGT 798
 |||
 Db 721 AAGTATATCAGCAAGCCGGAGAACCTG-AACTCATGATGAACCTNCTTCGGGAT-AAAGT 778
 QY 799 CCAACATCCAGTTTGAAGCCTTTCATGTT 828
 |||
 Db 779 CCAACATCCAGTTTGAAGCCTTCTGGTTTT 808

RESULT 5

AAX39817

ID AAX39817 standard; DNA; 831 BP.

XX

AC AAX39817;

XX

DT 02-JUL-1999 (first entry)

XX

DE Gastric cancer associated gene.

XX

KW Cancer associated antigen; diagnosis; research; treatment; human;
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 KW prostate cancer; ss.

XX

OS Homo sapiens.

XX

PN W09904265-A2.

XX

PD 28-JAN-1999.

KX
 PR 22-JUN-1998; 98US-0102322.
 PR 17-JUL-1997; 97US-0896164.
 PR 10-OCT-1997; 97US-0061599.
 PR 10-OCT-1997; 97US-0061765.
 PR 10-OCT-1997; 97US-0948705.
 PR 11-OCT-1997; 97GB-0021697.

XX
PI Chen Y, Gout I, Gure A, O'Hare M, Obata Y, Old LJ;
PI Pfreundschuh M, Sahin U, Scanlan MJ, Stockert E;
PI Tureci O;

XX New isolated cancer associated nucleic acids and polypeptides -
PT isolated using sera from cancer patients, used to develop products
PT for the diagnosis, monitoring or treatment of cancers

The invention relates to a method for diagnosing a disorder characterised by expression of a human cancer associated antigen precursor coded for by a nucleic acid molecule (NAM). The method comprises: (a) contacting a biological sample isolated from a subject with an agent that specifically binds to the NAM, an expression product or a fragment of an expression product complexed with an HLA molecule; and (b) determining the interaction between the agent and the NAM or the expression product as a determination of the disorder. The products and methods can be used in the diagnosis, monitoring, research, or treatment of conditions characterised by the expression of various cancer associated antigens. The invention provides nucleic acid sequences and encoded polypeptides which are cancer associated antigen precursors expressed in human breast cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and lung cancer.

Query Match 57.5%; Score 684.6; DB 20; Length 831;
Best Local Similarity 96.1%; Pred. No. 1.9e-180;
Matches 764; Conservative 0; Mismatches 23; Indels 8; Gaps 6;

Qy	1 ATGAAAAAATGCCTTTGTTTTAGTAAATCACACAAAATCCAGCAGAAATTGTGAAAATC	60
Db	37 ATGAAAAAATGCCTTTGTTTTAGTAAATCACACAAAATCCAGCAGAAATTGTGAAAATC	96
 Qy	 61 CTGAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA	 120
Db	97 CTGAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA	156
 Qy	 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	 180
Db	157 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA	216

Qy 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
 |||
 Db 217 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTG 276
 Qy 241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 |||
 Db 277 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 336
 Qy 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 |||
 Db 337 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 396
 Qy 361 GCTCATCCTCATATCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 |||
 Db 397 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 456
 Qy 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
 |||
 Db 457 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 516
 Qy 481 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||
 Db 517 TTTTCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 576
 Qy 547 TCAGATGCCTTTTGCTACTTTCAA-GGATTTACTAACCAGACATAAAGTGTGGTAGC-AG 598
 |||
 Db 577 TCAGATGCCTTTTGCTACTTTCAAGGGATTACTAACCAGACATAAAGTGTGGTAGCAAG 636
 Qy 607 ACTTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTG 658
 |||
 Db 637 ACTTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTG 696
 Qy 667 AG-AATTATGTTACTAAGAGACAGTCCTTAAAG-CTGCTAGGGGAGCTGATCCTGGACCG 716
 |||
 Db 697 AGAAATTATGTTACCAAGAGACAGTCCTTAAAGCCTGCTAAGGGAAGTATTCTGGACCG 756
 Qy 717 TCACAACTTTGCCATC-ATGACAAAGTATATCAGCAAGCC---GGAGAACCTGAAACTCA 772
 |||
 Db 757 TCANAACTTTGCCATCAANGCAAAAGTTTATCAACAAGCCNNGGGAAACCGGAAACNCA 816
 Qy 773 TGATGAACCTCCTTC 787
 |||
 Db 817 GGAGGAACCTCCTTC 831

RESULT 7

AAI60020/c

ID AAI60020 standard; cDNA; 1191 BP.

XX

AC AAI60020;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 4009.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;

KW

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Yang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 PI WPI; 2001-442253/47.
 PI P-PSDB; AAM40864.
 XX
 PI Novel nucleic acids and polypeptides, useful for treating disorders,
 PI such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 4009; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 1191 BP; 348 A; 261 C; 236 G; 346 T; 0 other;

Query Match 67.5%; Score 684.4; DB 22; Length 1191;
 Best Local Similarity 99.9%; Pred. No. 2.6e-180;

Matches 688; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      329 GCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCC 388
      |||||
Db      1189 GCACTCGAAGTCCTACTGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCC 1130

Qy      389 TCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTA 448
      |||||
Db      1129 TCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTA 1070

Qy      449 TTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTCTTTA 508
      |||||
Db      1069 TTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTCTTTA 1010

Qy      509 AGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATT 568
      |||||
Db      1009 AGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATT 950

Qy      569 TACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTT 628
      |||||
Db      949 TACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTT 890

Qy      629 TTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAA 688
      |||||
Db      639 TTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAA 839

Qy      689 AGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTGCCATCATGACAAAGTATATCA 748
      |||||
Db      639 AGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTGCCATCATGACAAAGTATATCA 770

Qy      749 GCAAGCCGAGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGTCCCAACATCC 808
      |||||
Db      759 GCAAGCCGAGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGTCCCAACATCC 710

Qy      809 AGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTA 868
      |||||
Db      769 AGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTA 650

Qy      869 TTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGCAGCTTCCAAA 928
      |||||
Db      649 TTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGCAGCTTCCAAA 590

Qy      929 AAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCC 988
      |||||
Db      539 AAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCC 530

Qy      989 GAGACTTGAAGAAAACGGCCCCCTTGA 1014
      |||||
Db      529 GAGACTTGAAGAAAACGGCCCCCTTGA 504
```

RESULT 8

AAC91772

ID AAC91772 standard; cDNA; 1026 BP.

XX

AC AAC91772;

XX

Qy	78	CATTTTGGAAAAGCAAGAC-----AAAAAGACAGACAAGGCTTCAGAAAGAAGTGTC	128
Db	72	TGTTCTGGAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC	131
Qy	129	TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAAC	188
Db	132	CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGAGCCTCAGAC	191
Qy	189	AGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTGACACTGAT	248
Db	192	AGAAGCAGTAGCTCAACTTGCTCAAGAAGTCTATAATAGTGGGCTCCTTAGCACCTGGT	251
Qy	249	AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAA	308
Db	252	AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGIGGCTCAAATTTTCAACAA	311
Qy	309	CATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCC	368
Db	312	TATTCTCAGAAGACAAATTGGTACGAGAACTCCTACTGTTGAATACATCTGCACCCAACA	371
Qy	369	TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG	428
Db	372	GAATATTTTGTTTCATGTTATTGAAAGGTATGAATCTCCAGAAATAGCTCAAATTTGTGG	431
Qy	429	GATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA	488
Db	432	AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA	491
Qy	489	TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC	548
Db	492	ACAGTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC	551
Qy	549	CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGA	608
Db	552	ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTGGGA	611
Qy	509	ACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT	668
Db	612	ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTACAGAAATATGT	671
Qy	669	TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC	728
Db	672	GACAAAAGACAGTCACTGAAGCTTCTCGGTGAACACTACTACTAGATAGACACAACTTCAC	731
Qy	729	CATCATGACAAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG	788
Db	732	AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG	791
Qy	789	GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTGTGGCCAG	848
Db	792	AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTTCACGTTTTTAAGGTGTTGTAGCCAA	851
Qy	849	TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACCTCATTGA	908
Db	852	TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACCTCATAGA	911

QY 909 GTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTTCGTGACGAGAAGAA 968
||| || ||| | | | | | | | | | | | | | | | | | | | | | | |
Db 912 GTTCCTCAGCAAGTTTCAGAACGACAGGACGCAGGATGAGCAGTTTAACGACGAGAAGAC 971

Qy 969 CTA CTATTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
||| || | | | | | | | | | | | | | | | | | | | | | | | | | |
Eb 972 CTATTTAGTTAAACAGATCAGGGATTTGAAGAGACCAGCTC 1012

RESULT 9

ABK13127

ID ABK13127 standard; cDNA; 3281 BP.

XX

AC ABK13127;

XX

DT 09-APR-2002 (first entry)

XX

XX
DE Human secretory polynucleotide (sptm) cDNA (481257.3).

XX

XX Signal peptide; transmembrane domain; human; sptm; ss; gene;
KW antiatherosclerotic; antipso

KW Signal peptide; transmembrane domain; human; Spem; 481257.3; antiarteriosclerotic; antiatherosclerotic; antipsoriatic;
KW 481257.3; antiarteriosclerotic; antiatherosclerotic; antidiabetic;

KW 481257.3; antiarteriosclerotic; antiatherosclerotic;
KW antiinflammatory; cytostatic; anti-HIV; antiallergic; antidiabetic;

antiinflammatory; cytostatic; analgesic; antitumor; neuroprotective; nephrotropic; antigout; antithyroid; hepatotropic; cancer;

KW osteopathic; antirheumatic; antiarthritic; dermatological; cancer;
KW nephrotropic; antigout; anticholesterolemia; antihypertensive;

KW immunosuppressive; antiulcer; ophthalmological; vulnerary; gout;
KW immunostimulative; antiparkinsonian;

KW immunosuppressive; antidiabetic; antiparkinsonian;
KW anticonvulsant; cerebroprotective; nootropic; arteriosclerosis

RW virucide; antibacterial; cell proliferative disorder; arteriosclerosis

KW atherosclerosis; psoriasis; immune system disorder; inflammation;
KW rheumatoid arthritis; diabetes mellitus; AIDS; Addison's disease;

EW required immunodeficiency syndrome; AIDS; Addison's disease;
EW anxiety distress syndrome; allergy; cirrhosis; osteoporosis

adult respiratory distress syndrome; allergy; cirrhosis; osteoporosis; diabetes mellitus; Graves' disease; multiple sclerosis; osteoarthritis

diabetes mellitus; Graves' disease; multiple sclerosis; osteoarthritis; rheumatoid arthritis; systemic lupus erythematosus; ulcerative colitis

rheumatoid arthritis; systemic lupus erythematosus; ulcerative colitis; hematopoietic cancer; neurological disorder; stroke; epilepsy;

haematopoietic cancer; neurological disorder; stroke; epilepsy;
Huntington's disease; Parkinson's disease; meningitis; prion disease

Huntington's disease; Parkinson's disease; meningitis; prion
disease; Creutzfeldt-Jakob disease; cerebral palsy; myasthenia gravis

diabetic neuropathy; Alzheimer's disease.

 $\gamma\bar{\chi}$

63 Homo sapiens.

XX

PN WO200111032-A1.

XX

PD 15-FEB-2001.

XX

01-JUN-2000; 2000WO-US15246.

- XX

PR 05-AUG-1999; 99US-147500P.

PR 05-AUG-1999; 99US-147501P.

xx

PA (INCY-) INCYTE GENOMICS INC.

XZ

PI Hodgson DM, Lincoln SE, Russo FD, Spiro AN, Cohen HI, Rosen BH, Chalup MS,

PI Bratcher SR, Dufour GE, Cohen HS, Rosen EM,
Gershenwald JB, Panzer SR, Roseberry AM;

PI Jones AL, Yu JY, Greenawalt LB, Farneser LL,
Diaz SE:

PI Wright RJ, Daniels SE,

X.

DR WPI; 2002-147236/19.

X

AC ABV22987;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 22978.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US05171.
 XX
 PR 17-FEB-2000; 2000US-183319P.
 PR 16-MAR-2000; 2000US-189862P.
 PR 25-MAY-2000; 2000US-207454P.
 PR 09-JUN-2000; 2000US-211314P.
 PR 18-JUL-2000; 2000US-219007P.
 PR 13-DEC-2000; 2000US-255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R. Endege WO, Monahan JE;
 XX
 IN WFI; 2001-662795/76.
 XX
 AB Novel isolated nucleic acid molecule associated with cancerous state of
 AB prostate cells and correlating with presence of prostate cancer, useful
 AB for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 EB Claim 1; Page 4088; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for:
 CC (a) assessing whether a patient is afflicted with prostate cancer;
 CC (b) monitoring the progression of prostate cancer in a patient;
 CC (c) assessing the efficacy of a test compound to inhibit prostate
 CC cancer in a patient;
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
 CC in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound;
 CC (g) determining whether prostate cancer has metastasized in a patient;
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a
 CC patient;
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
 XX
 SQ Sequence 3849 BP; 1142 A; 745 C; 858 G; 1081 T; 23 other;

Query Match 57.5%; Score 582.6; DB 23; Length 3849;
 Best Local Similarity 74.7%; Pred. No. 1e-151;
 Matches 748; Conservative 0; Mismatches 244; Indels 9; Gaps 1;

QY 18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAACACAATTTGGC 77
 ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 437 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 496

QY 78 CATTTTGAAAAAGCAAGAC-----AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 497 TGTCTGAAAAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 556

QY 129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAAC 188
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 557 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAATGAAAAGAGCCTCAGAC 616

QY 189 AGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 617 AGAAGCAGTAGCTCAACTTGCTCAAGAAGTCTATAATAGTGGGCTCCTTAGCACCTGGT 576

QY 249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAA 308
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 577 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTTCAACAA 536

QY 309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCTACTGTGGAGTATATTAGTGCTCATCC 368
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 737 TATTCTCAGAAGACAAATTGGTACGAGAACTCTACTGTTGAATACATCTGCACCCAACA 796

QY 809 TCAATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 128
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 857 GAATATTTTGTTCATGTTATTGAAAGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 856

QY 909 GATTATGCTGAGAGAATGTATCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 977 AATAATGTTAAGAGAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 916

QY 1039 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 917 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 976

QY 1149 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGTTAGCAGACTTCTTAGA 608
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 977 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTTTIGGA 1036

QY 1209 ACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1037 ACAGCATTATGATAGATTTTTTCAAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 1096

QY 1269 TACTAAGAGACAGTCTTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTTGC 728
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1097 GACAAAAGACAGTCACTGAAGCTTCTCGGTGAACCTACTACTAGATAGACACAACCTTCAC 1156

QY 1329 CATCATGACAAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1157 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1216

QY 1389 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
 || | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1217 AGACAAAAGTCGCAACATCCAGTTTGAAGCCTTTTACGTTTTTAAGGTGTTTGTAGCCAA 1276

[illegible]

RESULT 11

ABV28822

ID ABV28822 standard; cDNA; 3849 BP.

XX

AC ABV28822;

XX

DT 16-SEP-2002 (first entry)

XX

DE Human prostate expression marker cDNA 28813.

XX

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XW pharmacogenomic marker; gene; ss.



OS Homo sapiens.

 $\chi\chi$

PN WO200160860-A2.

$$-X$$

23-AUG-2001.

1

20-FEB-2001; 2001WO-US05171.

20

17-FEB-2000; 2000US-183319P.

16-MAR-2000; 2000US-189862P.

DR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

UR 13-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX

PT Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer -

XX

PS Claim 1; Page 6066-6067; 11750pp; English.

XZ

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;

QY 609 ACAAATACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGT 668
 Db 1037 ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTGAGAAATTATGT 1096
 QY 669 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
 Db 1097 GACAAAAGACAGTCACTGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACCTTCAC 1156
 QY 729 CATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
 Db 1157 AATTATGACAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 1216
 QY 789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGCGCCAG 848
 Db 1217 AGACAAAAGTCGCAACATCCAGTTTGAAGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 1276
 QY 849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAAACCTCATTGA 908
 Db 1277 TCCTAACAAGACGCAGCCCATCTAGACATCCTCCTCAAGAACCAGGCCAAACCTCATAGA 1336
 QY 909 GTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGACGAGAAGAA 968
 Db 1337 GTTCCTCAGCAAGTTTCAGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGAC 1396
 QY 969 CTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCC 1009
 Db 1397 CTATTTAGTTAAACAGATCAGGGATTGGAAGAGACCAGCTC 1437

RESULT 12

AAF30688

ID AAF30688 standard; cDNA; 1053 BP.

XX

AC

AAF30688;

XX

DT

11-JUN-2001 (first entry)

XX

DE

Human acute neuronal induced calcium binding protein ANIC-BP-1B cDNA.

XX

KW

Acute neuronal induced calcium binding protein; ANIC-BP-1B;
 splice variant; human; stroke; head trauma; Parkinson's disease;
 Alzheimer's disease; multiple sclerosis; spinal cord injury;
 cerebroprotective; antiparkinsonian; nootropic; neuroprotective;
 therapy; diagnosis; vaccine; ss.

XX

OS

Homo sapiens.

XX

PH

Key

Location/Qualifiers

FT

CDS

1..1053

FT

/*tag= a

FT

/product= "Human ANIC-BP-1B"

XX

PN

WO200125423-A1.

XX

PD

12-APR-2001.

XX

Db 192 AGAAGCAGTAGCTCAACTTGCTCAAGAACTCTATAATAGTGGGCTCCTTAGCACCCCTGGT 251

QY 249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAA 308
||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 252 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAGACGTGGCTCAAATTTTCAACAA 311

QY 309 CATCTTGAGAAGACAGATAGGCACTCGGAGTCTTACTGTGGAGTATATTAGTGCTCATCC 368
|| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 312 TATTCTCAGAAGACAAATTGGTACGAGAATCCTACTGTTGAATACATCTGCACCCAACA 371

QY 369 TCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGG 428
|||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 372 GAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATTGTGG 431

QY 429 GATTATGCTGAGAGAAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAA 488
|| ||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 432 AATAATGTTAAGAGAAATGCATCAGACATGAACCACTTGCAAAAATCATTTTGTGGTCGGA 491

QY 489 TCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGC 548
|| || | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 492 ACAGTTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCAGATGC 551

QY 549 CTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGTTAGTAGCAGACTTCTTAGA 608
|||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 552 ATTTGCCACATTCAAGGATTTACTTACAAGACATAAATGCTCAGTGCAGAAATTTTGGGA 611

QY 569 ACAAATACGACACTATTTTTGAAGACTATGAGAAATGCTTCAGTCTGAGAATTATGT 668
|| | |||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 572 ACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTCAGAAAATTATGT 674

QY 569 TACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGC 728
|| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 572 GACAAAAAGACAGTCACTGAAGCTTCTCGGTGAACCTACTACTAGATAGACACAACCTCAC 731

QY 729 CATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATGAACCTCCTTCG 788
|| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 732 AATTATGACAAAATACATCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCG 791

QY 789 GGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAG 848
|| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 792 AGACAAAAGTCGCAACATCCAGTTTGAGGCCTTTCACGTTTTTAAGGTGTTTGTAGCCAA 851

QY 849 TCCTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCCAACTCATTGA 908
|||| | |||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 852 TCCTAACAAGACGCAGCCCATCCTAGACATCCTCCTCAAGAACCAGGCCAACTCATAGA 911

QY 909 GTTTCTGAGCAGCTTCCAAAAGAAAGGACGGAT--GATGAGCAGTTCGCTGACGAGAAG 966
|| | |||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | |||||

Db 912 GTTCCTCAGCAAGTTTTCAGAACGACAGGACGGATTGTATGAGCAGTTCCGTACCGACGAC 971

QY 967 AACTAC 972
|| |

Db 972 GAATTC 977

ID AAS89557 standard; cDNA; 1162 BP.
XX
AC AAS89557;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #25361.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-GCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
JR WPI; 2001-639362/73.
JR P-PSDB; ABQ25370.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 25361; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1162 BP; 383 A; 241 C; 258 G; 280 T; 0 other;

Query Match 53.2%; Score 539.6; DB 23; Length 1162;
Best Local Similarity 73.9%; Pred. No. 5.6e-140;
Matches 743; Conservative 0; Mismatches 249; Indels 14; Gaps 4;

QY 18 GTTTAGTAAATCACACAAAAATCCAGCAGAAATGTGTGAAAATCCTGAAAAGACAATTTGGC 77
Db 143 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 202
QY 73 CATTTTGGAAAAGCAAGAC-----AAAAAGACAGACAAGGCTTCAGAAGAAGTGTG 128
Db 203 TGTCTGGAAGCAAGACATTTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 262
QY 129 TAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAAC 188
Db 263 CAAAAATCTGGTTGCCATGAAAGAAATTCTGTATGGCACAAATGAAAAAGATCCTCAGAC 322
QY 189 AGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCTGCTAGTGACACTGAT 248
Db 323 AGAAGCAGGAGCTCAACTTGCTCAAGAAGTCTATAATAGTGGGCTCCTTATCACCTGGT 382
QY 249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACA 308
Db 323 AGCTGATTTACAGCTCAATTGACTTTGAGGGCAAAAAAGACGTGGCTCAAATTTCAACA 442
QY 349 CATCTTGAGAAGACAGATAGGCA-CTCGAGTCTACTGTGGAGTATATTAGTGCTCATC 367
Db 443 TATTCTCAGAAGACAAATTGGTACCGAGAACTCCTACTGTTGAATACATCTGCACCCAAA 502
QY 468 CTCA--TATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATT--GCCTTACGT 423
Db 503 CAGAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATT 562
QY 484 TGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTT 483
Db 563 TGTGGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGGCAAAATCATTTTGTGG 522
QY 484 TCTAATCAATTCAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCA 543
Db 523 TCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCA 682
QY 544 GATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTC 603
Db 693 GATGCATTTGCCACATTCAAGGATTTACTTACAAGACATAAATTGCTCAGTGCAGAATTT 742
QY 604 TTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAAT 663
Db 743 TTGGAACAGCATTATGATAGATTTTTCAGTGAATATGAGAAGTTACTTCATTTCAGAAAAT 802
QY 664 TATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAAC 723
Db 803 TATGTGACAAAAGACAGTCACTGAAGCTTCTCGGTGAAGTACTACTAGATAGACACAAC 862
QY 724 TTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACC'TGAAACTCATGATGAACCTC 783

[illegible]

RESULT 14

ID: AAX39818 standard; DNA; 833 BP.

AC

AC AAX39818;

ST

3-17

100

100

100

2007

567

23

250
251

2000

05

XIV

PN

3X

תק

20
557 Δ

P5

XX

De

DT

52

51
52
53

21.

10

PH

XX

PD

75
76

227

2.

PE

p2

X2

DE

21
22

△△
△△

D.

XX
PS Claim 67; Page 559; 787pp; English.

Sequence 833 BP; 253 A; 171 C; 172 G; 227 T; 10 other;

Query Match 51.3%; Score 520.2; DB 20; Length 833;
Best Local Similarity 98.1%; Pred. No. 1.2e-134;
Matches 566; Conservative 0; Mismatches 7; Indels 4; Gaps 4;

Qy	442	GAATGTATTTCGACATGAACCACTTG-CCAAAATCATCCTC-TTTTCTAATCAATTTCAGAG	499
Db	732	GAATNTATTTCGACTTGACCCANTTGCCCAAATCATCCTCTTTTCTAATCAATTTCAGAG	673
Qy	500	ATTTCTTTTAAGT-ACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT	558
Db	672	ATTTCTTTTAAGTAACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACT	613
Qy	559	TTCAAGGATTTTACTAACCAGA-CATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTA	617
Db	612	TTCAAGGATTTTACTAACCNGACCTTAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTA	553
Qy	518	CGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAG	677
Db	552	CGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAG	493
Qy	678	ACAGTCTTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTTGCCATCATGAC	737
Db	592	ACAGTCTTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTTGCCATCATGAC	433
Qy	738	AAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAG	797
Db	432	AAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAG	373
Qy	798	TCCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAA	857
Db	372	TCCCAACATCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAA	313
Qy	858	AACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAG	917
Db	312	AACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAG	253

QY 918 CAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGACGAGAAGAACTACTTGAT 977
 |||||
 Db 252 CAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGACGAGAAGAACTACTTGAT 193
 QY 978 TAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
 |||||
 Db 192 TAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 156

RESULT 15

AAS88031

ID AAS88031 standard; cDNA; 2492 BP.

XX

AC AAS88031;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #23835.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

EN NM_00175067-A2.

XX

DD 21-OCT-2001.

XX

PP 01-MAR-2001; 2001WO-US08631.

XX

PP 01-MAR-2000; 2000US-0540217.

XX

PP 03-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drymanac RT; Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

XX

DR P-PSDB; ABG23844.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX

PS Claim 1; SEQ ID No 23835; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 2492 BP; 751 A; 477 C; 546 G; 718 T; 0 other;

Query Match 48.9%; Score 496; DB 23; Length 2492;
 Best Local Similarity 73.0%; Pred. No. 1.1e-127;
 Matches 737; Conservative 0; Mismatches 255; Indels 18; Gaps 7;

QY 19 GTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAATCCTGAAAGACAATTTGGC 77
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 143 GTTTGGGAAGTCTCACAAATCTCCAGCAGACATTGTGAAGAATCTGAAGGAGAGCATGGC 202
 QY 78 CATTTTGGAAAAGCAAGAC-----AAAAAGACAGACAAGGCTTCAGAAGAAGTGTC 128
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 203 TGTTCCTGGAAGCAAGACATTCTGATAAAAAAGCAGAAAAGGCTACAGAAGAAGTTTC 262
 QY 339 TAAATCACTGCAAGCAATGAAAGAAATTCGTGTGGTACAAACGAGAAAAGAACCCCAAC 188
 ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 323 CAAAAATCTGGTTGCCATGAAAGAAATTCGTATGGCACAAATGAAAAAGATCCTCAGAC 322
 QY 359 AGAAGCAGTGGCTCAGCTAGCACAGAAGTCTACAGCAGTGGCCTGCTAGTGACACTGAT 248
 ||||| ||||| ||||| | | | | | | | | | | | | | | | | | | | | | | | |
 Db 323 AGAAGCAGCAGCTCAACTTGCTCAAGAAGTCTATAATAGTGGGCTCCTTATCACCTGGT 382
 QY 249 AGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAA 308
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 383 AGCTGATTTACAGCTCATTGACTTTGAGGGCAAAAAGACGTGGCTCAAATTTTCAACAA 442
 QY 309 CATCTTGAGAAGACAGATAGGCA-CTCGGAGTCTACTGTGGAGTATATTAGTGCTCATC 367
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 443 TATTCTCAGAAGACAAATTGGTACCGAGAAGTCTACTGTTGAATACATCTGCACCCAAA 502
 QY 358 CTCA--TATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATT--GCCTTACGT 423
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 503 CAGAATATTTTGTTCATGTTATTGAAAGGGTATGAATCTCCAGAAATAGCTCTAAATT 562
 QY 424 TGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTT 483
 ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 563 TGTGGAATAATGTTAAGAGAATGCATCAGACATGAACCACTTGGCAAAATCATTTTGTGG 622
 QY 434 TCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCA 543
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 623 TCGGAACAGTTTATGATTTCTTCAGATATGTCGAAATGTCAACATTTGACATAGCTTCA 682
 QY 544 GATGCC-TTGTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTT 602
 ||||| ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 01:16:12 ; Search time 94 Seconds
(without alignments)
4761.303 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcctttgtt.....tgaagaaaacggccccttga 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					Description
Result	Query	Match	Length	DB	ID		
No.	Score						
1	1014	100.0	1421	4	US-09-620-312D-111		Sequence 111, App
2	1010.8	99.7	1344	3	US-09-190-965-2		Sequence 2, Appli
3	1010.8	99.7	1344	4	US-09-470-253-2		Sequence 2, Appli
c 4	51.6	5.1	7218	1	US-08-232-463-14		Sequence 14, Appl
5	37.8	3.7	1457	4	US-09-214-307A-9		Sequence 9, Appli
c 6	37.4	3.7	4103	4	US-09-620-312D-390		Sequence 390, App
7	36.4	3.6	4533	3	US-08-726-214-5		Sequence 5, Appli
c 8	35.6	3.5	577	4	US-09-513-057C-20		Sequence 20, Appl
c 9	35.6	3.5	3707	3	US-09-276-531-42		Sequence 42, Appl
c 10	35.6	3.5	5714	4	US-09-620-312D-393		Sequence 393, App
11	35.2	3.5	12793	4	US-09-004-838-124		Sequence 124, App

	12	35	3.5	775	3	US-08-961-083-89	Sequence 89, Appl
	13	35	3.5	775	4	US-09-536-784-89	Sequence 89, Appl
	14	35	3.5	1696	3	US-08-961-083-217	Sequence 217, App
	15	35	3.5	1696	4	US-09-536-784-217	Sequence 217, Appl
	16	35	3.5	3763	1	US-07-792-865D-1	Sequence 1, Appli
c	17	35	3.5	13926	4	US-08-961-527-5	Sequence 5, Appli
c	18	34.8	3.4	1413	4	US-09-016-434-1382	Sequence 1382, Ap
	19	34.8	3.4	12734	4	US-09-344-456-1	Sequence 1, Appli
	20	34.8	3.4	14078	3	US-09-433-262-1	Sequence 1, Appli
	21	34.8	3.4	14078	4	US-09-702-330-1	Sequence 1, Appli
	22	34.8	3.4	14578	3	US-08-859-694-1	Sequence 1, Appli
c	23	34.6	3.4	64467	4	US-09-803-671B-3	Sequence 3, Appli
	24	34.4	3.4	22481	4	US-08-367-841A-43	Sequence 43, Appl
	25	34.4	3.4	22481	5	PCT-US95-07201-43	Sequence 43, Appl
	26	34.4	3.4	22484	4	US-09-875-223-2	Sequence 2, Appli
	27	34	3.4	3378	4	US-09-328-352-4107	Sequence 4107, Ap
c	28	34	3.4	43360	4	US-09-453-702B-206	Sequence 206, App
c	29	34	3.4	45325	4	US-09-453-702B-261	Sequence 261, App
	30	33.8	3.3	189	4	US-09-134-001C-1145	Sequence 1145, Ap
c	31	33.8	3.3	4233	3	US-09-056-105-27	Sequence 27, Appl
	32	33.6	3.3	1276	3	US-09-177-325-2	Sequence 2, Appli
	33	33.6	3.3	1276	3	US-09-411-812A-2	Sequence 2, Appli
	34	33.6	3.3	1276	4	US-09-590-113-2	Sequence 2, Appli
	35	33.6	3.3	1347	4	US-09-134-001C-1309	Sequence 1309, Ap
	36	33.4	3.3	1410	4	US-09-328-352-124	Sequence 124, App
c	37	33.4	3.3	1508	1	US-08-236-311-5	Sequence 5, Appli
c	38	33.4	3.3	1508	3	US-08-457-918-5	Sequence 5, Appli
	39	33.4	3.3	2643	4	US-09-486-072-6	Sequence 6, Appli
c	40	33.4	3.3	18627	4	US-08-961-527-113	Sequence 113, App
	41	33.2	3.3	378	4	US-09-252-991A-10313	Sequence 10313, A
c	42	33.2	3.3	813	4	US-09-252-991A-10547	Sequence 10547, A
	43	33.2	3.3	831	4	US-09-252-991A-10512	Sequence 10512, A
c	44	33.2	3.3	3253	4	US-09-333-214-4	Sequence 4, Appli
	45	33.2	3.3	8091	4	US-09-230-652-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-620-312D-111

; Sequence 111, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunqing


```
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 111
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (217)..(1230)
US-09-620-312D-111
```

```
Query Match          100.0%; Score 1014; DB 4; Length 1421;
Best Local Similarity 100.0%; Pred. No. 4.8e-292;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
        |||
Db      217 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276

Qy      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
        |||
Db      277 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 336

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
        |||
Db      337 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 396

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
        |||
Db      397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTG 456

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
        |||
Db      457 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 516

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
        |||
Db      517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
        |||
Db      577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636

Qy      421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
```



```

      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      637 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 696

Qy      481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      697 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 756

Qy      541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAAGTGTGGTAGCAGAC 600
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      757 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAAGTGTGGTAGCAGAC 816

Qy      601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      817 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 876

Qy      661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      877 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 936

Qy      721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      937 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 996

Qy      781 CTCCTTCGGGATAAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      997 CTCCTTCGGGATAAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 1056

Qy      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 900
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1057 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 1116

Qy      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGAC 960
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1117 CTCATTGAGTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTTCGCTGAC 1176

Qy      961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1177 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1230

```

RESULT 2

US-09-190-965-2

; Sequence 2, Application US/09190965

; Patent No. 6071721

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/09/190,965

; CURRENT FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program

; SEQ ID NO 2

; LENGTH: 1344


```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-09-190-965-2
```

```
Query Match          99.7%; Score 1010.8; DB 3; Length 1344;
Best Local Similarity 99.8%; Pred. No. 4.2e-291;
Matches 1012; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy      1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 183

Qy      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      184 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
          ||||| ||||||||||||||||||||||||||||||||||||||||| |||
Db      304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTGGTG 363

Qy      241 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      364 ACACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 423

Qy      301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

Qy      421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      544 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 603

Qy      481 TTTTCTAATCAATTGAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      604 TTTTCTAATCAATTGAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

Qy      541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 723

Qy      601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy      661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      784 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843
```



```

Qy      721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
        |||
Db      844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy      781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
        |||
Db      904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

Qy      841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAAA 900
        |||
Db      964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAAA 1023

Qy      901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
        |||
Db     1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

Qy      961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
        |||
Db     1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1137

```

RESULT 3

US-09-470-253-2

; Sequence 2, Application US/09470253

; Patent No. 6365371

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/09/470,253

; CURRENT FILING DATE: 1999-12-22

; PRIOR APPLICATION NUMBER: 09/190,965

; PRIOR FILING DATE: 1998-11-13

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PERL Program

; SEQ ID NO 2

; LENGTH: 1344

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: 3734805

US-09-470-253-2

```

Query Match          99.7%;  Score 1010.8;  DB 4;  Length 1344;
Best Local Similarity 99.8%;  Pred. No. 4.2e-291;
Matches 1012;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

```

```

Qy      1 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
        |||
Db     124 ATGAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 183

Qy      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
        |||

```


Db 184 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 243

Qy 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
 |||||

Db 244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

Qy 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
 |||||

Db 304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTGGTG 363

Qy 241 ACACCTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 |||||

Db 364 ACACCTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 423

Qy 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 |||||

Db 424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 483

Qy 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 |||||

Db 484 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

Qy 421 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 480
 |||||

Db 544 CGTTGTGGGATTATGCTGAGAGAATGTATTGACATGAACCACTTGCCAAAATCATCCTC 603

Qy 481 TTTTCTAATCAATTCAAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||||

Db 604 TTTTCTAATCAATTCAAGAGATTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

Qy 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 600
 |||||

Db 664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGAC 723

Qy 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 |||||

Db 724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy 661 AATTATGTTACTAAGAGACAGTCTTTAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 |||||

Db 784 AATTATGTTACTAAGAGACAGTCTTTAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843

Qy 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
 |||||

Db 844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 |||||

Db 904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 963

Qy 841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCCAAA 900
 |||||

Db 964 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCCAAA 1023

Qy 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 |||||

Db 1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

Qy 961 GAGAAGAACTACTTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 |||||
 Db 1084 GAGAAGAACTACTTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1137

RESULT 4

US-08-232-463-14/c

; Sequence 14, Application US/08232463
 ; Patent No. 5670367
 ; GENERAL INFORMATION:
 ; APPLICANT: DORNER, F.
 ; APPLICANT: SCHEIFLINGER, F.
 ; APPLICANT: FALKNER, F. G.
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 1800 Diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,463
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/935,313
 ; FILING DATE:
 ; APPLICATION NUMBER: EP 91 114 300.6
 ; FILING DATE: 26-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703)836-9300
 ; TELEFAX: (703)683-4109
 ; TELEX: 899149
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7218 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; CLONE: pTZgpt-Fls
 US-08-232-463-14

Query Match 5.1%; Score 51.6; DB 1; Length 7218;
 Best Local Similarity 3.6%; Pred. No. 4.2e-05;

Qy 490 CAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCC 549
 ||| | | | | | | | | | | | | | | |
 Db 446 TTATTTTATCAATTTAGTAAGGATGCCAACTTTGAACCTATTGCTTGTAGACCCTATCGT 505
 Qy 550 TTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGACTTCTTAGAA 609
 | | | | | | | | | | | | | | | | | |
 Db 506 CCTCAAACAAAAGGGTCTGTTGAATCATTAGCTAAATTTGTTGAACAGCGTTTAAGACCA 565
 Qy 610 CAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTT 669
 | | | | | | | | | | | | | | | | | |
 Db 566 TACGATTATGAATTTTATGATGCTG--TAGAACTTATTGGGCTAGTAAACGATTATGTC 623
 Qy 670 ACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCC 729
 || || | | | | | | | | | | | | | | | |
 Db 624 ACGAATTGAATCACTTAGAAATTTACAAGCAACAGAACACGACCTATCGACGTTTTC 683
 Qy 730 ATCATGACAAAGTATATCA 748
 || || | | | | | | | | | |
 Db 684 ATTATGAAGAAAAGAACA 702

RESULT 6

US-09-620-312D-390/c

; Sequence 390, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunqing

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: John Tillinghast

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 390

; LENGTH: 4103


```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)..(3493)
US-09-620-312D-390
```

```
Query Match          3.7%; Score 37.4; DB 4; Length 4103;
Best Local Similarity 60.2%; Pred. No. 0.53;
Matches 62; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
```

```
Qy      6 AAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAA 65
        | | | | |      | | | | | | | | | | | | | | | | | | | | | |
Db      4091 ACAAATGAGAAAGTTTCATTTACCTCAAAAAAATCCAGGCTATACAAACAGACAACCTGAA 4032

Qy      66 AGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGAC 108
        | | | | | | | | | | | | | | | | | | | | | |
Db      4031 AGCCACATAGGAAATTTCCGAAACACAAAAGAAAAAGTCTCAC 3989
```

RESULT 7

US-08-726-214-5

```
; Sequence 5, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,214
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:450
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 5:
```



```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4533 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-726-214-5
```

```
Query Match          3.6%; Score 36.4; DB 3; Length 4533;
Best Local Similarity 56.8%; Pred. No. 1.1;
Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
```

```
Qy      718 CACAACCTTTGCCATCATGACAAAGTATATCAGCAAGCCGAGAACCTGAAACTCATGATG 777
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2644 CTCATCGCCACCATCATGCTGGTGCAGGTCAGCCACATGGTGAAGCTGACACTCATGCTG 2703

Qy      778 AACCTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGG 835
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2704 CTCGTCACAGGCGCCGTGACTGCCATCAACCTGTATGCCTGGTGTCTCTTTGATG 2761
```

RESULT 8

```
US-09-513-057C-20/c
; Sequence 20, Application US/09513057C
; Patent No. 6433251
; GENERAL INFORMATION:
; APPLICANT: Wagner, et al.
; TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND
PHOTOPERIODISM
; FILE REFERENCE: 1505-54357
; CURRENT APPLICATION NUMBER: US/09/513,057C
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 577
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-09-513-057C-20
```

```
Query Match          3.5%; Score 35.6; DB 4; Length 577;
Best Local Similarity 51.2%; Pred. No. 0.64;
Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

```
Qy      457 GAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTCTTTAAGTACGTG 516
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      223 GACCCAAATACCCAAAACACAATCTTTACATAGAAATCAAGAGATTCTGAAGCACACAG 164

Qy      517 GAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACC 576
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      163 AAGCAAAAAGATGTATAATTTACAAAATTACTATTATATTTTCTGTGATCATGTAAC 104

Qy      577 AGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAATTAC 618
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      103 AGGCCTTGTTGGTAAGCACATAATATGAAGAAAGAGATTAC 62
```

RESULT 9

US-09-276-531-42/c
 ; Sequence 42, Application US/09276531
 ; Patent No. 6183968
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Reddy, Roopa
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Baughn, Mariah R.
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
 ; TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL
 PROLIFERATION
 ; NUMBER OF SEQUENCES: 134
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/276,531
 ; FILING DATE: Herewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/079,677
 ; FILING DATE: March 27, 1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lynn E. Murry, Ph.D.
 ; REGISTRATION NUMBER: 42,918
 ; REFERENCE/DOCKET NUMBER: PA-0008 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3707 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: CERVNOT01
 ; CLONE: 936117
 US-09-276-531-42

Query Match 3.5%; Score 35.6; DB 3; Length 3707;
 Best Local Similarity 51.9%; Pred. No. 1.7;
 Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

Qy 445 TGTATTGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTC 504
 || ||| | ||| || | |||| | ||||| || || | || ||
 Db 3154 TGCTTTCAAAATGTGGAACAACTAAAATATAAGGCTTTTCTGATAAACTATAAAAATTT 3095
 Qy 505 TTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAG 564
 | || || |||| | |||| | | | | | | ||| ||||
 Db 3094 AATCAGCACTTGGATCTAATGACATATCTTTATAATACTTCCTCTGCAGATACATTCACT 3035
 Qy 565 GATTTACTAACCAGACATAAAGTGTGGTAGCAG 598
 | || | | |||| | ||| || ||
 Db 3034 TAGTTCAAACCTTAACATACAAAGTTAGTCTCAG 3001

RESULT 10

US-09-620-312D-393/c

; Sequence 393, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunqing

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: John Tillinghast

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 393

; LENGTH: 5714

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (272)..(4312)

US-09-620-312D-393

Query Match 3.5%; Score 35.6; DB 4; Length 5714;
 Best Local Similarity 51.9%; Pred. No. 2.2;

Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```

Qy      445 TGTATTGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAGAGATTTC 504
      || ||| | ||| || | |||| | ||||| || || | | |||
Db      5233 TGCTTTCAAAATGTGGAACAACTAAAAATATAAGGCTTTTCTGATAAACTATAAAAATTT 5174

Qy      505 TTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAG 564
      | || || |||| | |||| | | | | | | | | ||| ||||
Db      5173 AATCAGCACTTGGATCTAATGACATATCTTTGTAATACTTCCTCTGCAGATACATTCACT 5114

Qy      565 GATTTACTAACCAGACATAAAGTGTTGTTAGTAGCAG 598
      | || | | |||| | ||| || |||
Db      5113 TAGTTCAAACCTTAACATACAAAGTTAGTCTCAG 5080

```

RESULT 11

US-09-004-838-124

; Sequence 124, Application US/09004838

; Patent No. 6350933

; GENERAL INFORMATION:

; APPLICANT: Michelmores, Richard W.

; APPLICANT: Shen, Kathy

; APPLICANT: Meyers, Blake

; TITLE OF INVENTION: Procedures and Materials for

; TITLE OF INVENTION: Conferring Pest Resistance in Plants

; NUMBER OF SEQUENCES: 140

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/004,838

; FILING DATE: 09-JAN-1998

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/781,734

; FILING DATE: 10-JAN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Einhorn, Gregory P.

; REGISTRATION NUMBER: 38,440

; REFERENCE/DOCKET NUMBER: 023070-078810US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 124:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12793 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single


```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..12793
; OTHER INFORMATION: /note= "RG2S"
US-09-004-838-124
```

```
Query Match          3.5%; Score 35.2; DB 4; Length 12793;
Best Local Similarity 47.6%; Pred. No. 4.4;
Matches 101; Conservative 10; Mismatches 98; Indels 3; Gaps 1;
```

```
Qy      438 GAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATCAATTCAG 497
          |||:| : : || |: :|:|:| | | | | | | | | | | |
Db      5998 GAGARAGWAWGRRRGAKAKARMCSMSYTTGGGATGTGATACTTCTTTTAGGAAAATGGAG 6057

Qy      498 AGATTTCTTTAAGTACGTGGAGTTGTCA---ACATTTGATATTGCTTCAGATGCCTTTGC 554
          || ||||| | || || | | | | | | | | | | | | |
Db      6058 TTATATCTTTGATATTGTATTTTTTTAATGTAATTTATATATTTAATCATTTTAGTTTAT 6117

Qy      555 TACTTTCAAGGATTTACTAACCAGACATAAAGTGTTGGTAGCAGACTTCTTAGAACAAAA 614
          | ||| | |||| | | | | |||| | | | | | | | | |
Db      6118 AAGTTTTATTTATTTTGATATGAAAAAAAAGTCTTTTATACATTGGATTTAACATAAAA 6177

Qy      615 TTACGACACTATTTTTGAAGACTATGAGAAAT 646
          | | ||| |||| | || | | | | | |
Db      6178 ATCCAACAATATTAATCAAAAAGACCAMACAT 6209
```

RESULT 12

US-08-961-083-89

```
; Sequence 89, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
```



```

; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-083-89

```

```

Query Match          3.5%; Score 35; DB 3; Length 775;
Best Local Similarity 46.9%; Pred. No. 1.1;
Matches 143; Conservative 0; Mismatches 160; Indels 2; Gaps 1;

```

```

Qy      22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
      ||| || || | || |||| || ||| | || | |
Db      263 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 322

Qy      82 TTGGAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
      | || |||| |||| || | |||| | | || | | ||
Db      323 GCGCCAAGAGAAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 382

Qy      142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTGGCT 201
      | | ||| | | | | || | ||| | | ||| | |||
Db      383 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAGTCAACTCCAGATACCAAGGCT 442

Qy      202 CAGCTAGCACAGAAGCTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
      | | | |||| | | | || |||| || | || | ||
Db      443 GAAGAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAAG 500

Qy      262 CTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
      ||| | | | | |||| | || ||| ||| | | ||| |
Db      501 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 560

Qy      322 CAGAT 326
      ||| |
Db      561 CAGGT 565

```

RESULT 13

US-09-536-784-89

```

; Sequence 89, Application US/09536784
; Patent No. 6573082

```

GENERAL INFORMATION:

```

; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA

```



```

;      ZIP: 20850
;
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER: HP Vectra 486/33
;      OPERATING SYSTEM: MSDOS version 6.2
;      SOFTWARE: ASCII Text
;
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/536,784
;      FILING DATE: 30-Oct-1997
;      CLASSIFICATION: <Unknown>
;
;      PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 08/961,083
;      FILING DATE: OCT-30-1997
;
;      ATTORNEY/AGENT INFORMATION:
;      NAME: Michelle S. Marks
;      REGISTRATION NUMBER: 41,971
;      REFERENCE/DOCKET NUMBER: PB340P3
;
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (301) 309-8504
;      TELEFAX: (301) 309-8512
;
;      INFORMATION FOR SEQ ID NO: 89:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 775 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;
;      SEQUENCE DESCRIPTION: SEQ ID NO: 89:
US-09-536-784-89

```

```

Query Match          3.5%;  Score 35;  DB 4;  Length 775;
Best Local Similarity 46.9%;  Pred. No. 1.1;
Matches 143;  Conservative 0;  Mismatches 160;  Indels 2;  Gaps 1;

```

```

Qy      22 AGTAAATCACACAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
      ||| || || | || |||| || || | || | || |
Db      263 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 322

Qy      82 TTGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
      | || |||| |||| || | |||| | | || || ||
Db      323 GCGCCAAGAGAAGACGAAAAGGCACCGTCTGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 382

Qy      142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCAACAGAAGCAGTGGCT 201
      | | ||| | | | | || | || | || | || | |||
Db      383 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 442

Qy      202 CAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
      | | | |||| | | | || | ||| || | | || | ||
Db      443 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAAG 500

Qy      262 CTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
      ||| | | | | |||| | || || | || | | ||| |
Db      501 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAGTTGAACAAG 560

Qy      322 CAGAT 326
      ||| |
Db      561 CAGGT 565

```


RESULT 14

US-08-961-083-217

; Sequence 217, Application US/08961083

; Patent No. 6159469

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,083

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 217:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1696 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-083-217

Query Match 3.5%; Score 35; DB 3; Length 1696;

Best Local Similarity 46.9%; Pred. No. 1.7;

Matches 143; Conservative 0; Mismatches 160; Indels 2; Gaps 1;

```

Qy      22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
      ||| || || | || |||| || || | || | || |
Db      275 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAGTTGAACAAGCAGGTGAACCAGTC 334

Qy      82 TTGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
      | || |||| |||| || | |||| | | || | ||
Db      335 GCGCCAAGAGAAGACGAAAAGGCACCGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 394

Qy     142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAACCCCCAACAGAAGCAGTGGCT 201
      | | ||| | | | | || | ||| | | ||| | |||

```


Db 395 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 454
 Qy 202 CAGCTAGCACAGAAGAACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
 | | | | | | | | | | | | | | | | | | | |
 Db 455 GAAGAAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAG 512
 Qy 262 CTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
 | | | | | | | | | | | | | | | | | | | |
 Db 513 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAAGTTGAACAAG 572
 Qy 322 CAGAT 326
 | | | |
 Db 573 CAGGT 577

RESULT 15

US-09-536-784-217

; Sequence 217, Application US/09536784

; Patent No. 6573082

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/536,784

; FILING DATE: 30-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB340P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 217:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1696 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 217:

US-09-536-784-217

Query Match 3.5%; Score 35; DB 4; Length 1696;
 Best Local Similarity 46.9%; Pred. No. 1.7;
 Matches 143; Conservative 0; Mismatches 160; Indels 2; Gaps 1;

```

Qy      22 AGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATCCTGAAAGACAATTTGGCCATT 81
      ||| || || | || |||| || ||| | || | |
Db      275 AGTCAACCATCAGACAAACCAGCTGAGGAATCAAAAAGTTGAACAAGCAGGTGAACCAGTC 334

Qy      82 TTGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAAGAAGTGTCTAAATCACTGCAA 141
      | || |||| |||| || | |||| | | || | | ||
Db      335 GCGCCAAGAGAAGACGAAAAGGCACCAGTCGAGCCAGAAAAGCAACCAGAAGCTCCTGAA 394

Qy     142 GCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAAGAACCCCCAACAGAAGCAGTGGCT 201
      | | ||| | | | | || | ||| | | ||| | |||
Db      395 GAAGAGAAGGCTGTAGAGGAAACACCGAAACAAGAAGAGTCAACTCCAGATACCAAGGCT 454

Qy     202 CAGCTAGCACAGAAGACTCTACAGCAGTGGCCTGCTAGTGACACTGATAGCTGACCTGCAG 261
      | | | |||| | | | || | ||| || | || | ||| ||
Db      455 GAAGAACTGTAGAA--CCAAAAGAGGAGACTGTTAATCAATCTATTGAACAACCAAAAAG 512

Qy     262 CTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATATTTAACAACATCTTGAGAAGA 321
      ||| | | | | |||| | || ||| ||| | | |||| |
Db      513 TTGAAACGCCTGCTGTAGAAAAACAAACAGAACCAACAGAGGAACCAAAAAGTTGAACAAG 572

Qy     322 CAGAT 326
      ||| |
Db      573 CAGGT 577

```

Search completed: January 6, 2004, 03:19:48
 Job time : 96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 02:35:04 ; Search time 1394 Seconds
(without alignments)
2517.743 Million cell updates/sec

Title: US-10-088-872-1
Perfect score: 1014
Sequence: 1 atgaaaaaaatgcstttgtt.....tgaagaaaacgccccttga 1014

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query					Description
No.	Score	Match Length	DB	ID			

1	1014	100.0	1421	13	US-10-117-722-111	Sequence 111, App
2	1014	100.0	1421	15	US-10-037-270-111	Sequence 111, App
3	1010.8	99.7	1344	15	US-10-025-730-2	Sequence 2, Appli
4	398	39.3	475	11	US-09-918-995-5343	Sequence 5343, Ap
5	288.8	28.5	690	9	US-09-910-943-318	Sequence 318, App
6	246.4	24.3	435	10	US-09-867-701-5263	Sequence 5263, Ap
7	244.8	24.1	447	10	US-09-867-701-5899	Sequence 5899, Ap
8	244.8	24.1	450	10	US-09-867-701-4953	Sequence 4953, Ap
9	210.8	20.8	762	9	US-09-910-943-35	Sequence 35, Appl
c 10	195	19.2	387	10	US-09-954-456-1453	Sequence 1453, Ap
c 11	195	19.2	387	10	US-09-880-107-481	Sequence 481, App
12	169.8	16.7	722	13	US-10-257-826A-118	Sequence 118, App
13	166.6	16.4	700	13	US-10-257-826A-119	Sequence 119, App
14	156	15.4	861	9	US-09-770-445-592	Sequence 592, App
15	74.2	7.3	262	9	US-09-923-876-1251	Sequence 1251, Ap
16	74.2	7.3	262	12	US-09-923-876-1251	Sequence 1251, Ap
c 17	65.6	6.5	336	11	US-09-918-995-19069	Sequence 19069, A
18	65	6.4	487	12	US-10-242-355-323	Sequence 323, App
19	65	6.4	487	13	US-10-080-254-54	Sequence 54, Appl
20	53.6	5.3	254	10	US-09-878-574-13369	Sequence 13369, A
c 21	50.6	5.0	486	11	US-09-770-961-777	Sequence 777, App
22	41.6	4.1	242	9	US-09-923-876-2528	Sequence 2528, Ap
23	41.6	4.1	242	12	US-09-923-876-2528	Sequence 2528, Ap
24	40.8	4.0	1295	12	US-10-310-154-294	Sequence 294, App
25	40.2	4.0	113306	12	US-10-292-798-1007	Sequence 1007, Ap
26	39.8	3.9	431	11	US-09-918-995-5787	Sequence 5787, Ap
c 27	38.4	3.8	6301	13	US-10-311-455-26	Sequence 26, Appl
28	38.2	3.8	1200	13	US-10-027-632-261235	Sequence 261235, A
29	38.2	3.8	1200	14	US-10-027-632-261235	Sequence 261235, A
30	37.8	3.7	1457	15	US-10-054-968-9	Sequence 9, Appli
31	37.8	3.7	7178	13	US-09-873-367C-278	Sequence 278, App
32	37.6	3.7	1267	14	US-10-601-843-45	Sequence 45, Appl
c 33	37.6	3.7	3673778	13	US-10-312-841-1	Sequence 1, Appli
34	37.4	3.7	2232	15	US-10-087-464-45	Sequence 45, Appl
c 35	37.4	3.7	4012	9	US-09-876-889-335	Sequence 335, App
c 36	37.4	3.7	4103	13	US-10-117-722-390	Sequence 390, App
c 37	37.4	3.7	4103	15	US-10-037-270-390	Sequence 390, App
c 38	37.4	3.7	8577	13	US-10-311-455-1760	Sequence 1760, Ap
c 39	37.2	3.7	869	13	US-10-027-632-261978	Sequence 261978, A
c 40	37.2	3.7	869	14	US-10-027-632-261978	Sequence 261978, A
c 41	37	3.6	5413	13	US-10-311-455-538	Sequence 538, App
c 42	36.6	3.6	9367	13	US-10-311-455-944	Sequence 944, App
43	36.4	3.6	461	14	US-10-079-623-143	Sequence 143, App
44	36.4	3.6	2641	12	US-10-369-493-29299	Sequence 29299, A
c 45	36.4	3.6	6071	13	US-10-311-455-297	Sequence 297, App

ALIGNMENTS

RESULT 1

US-10-117-722-111

; Sequence 111, Application US/10117722

; Publication No. US20030219744A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom


```

; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 111
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (217)..(1230)
US-10-117-722-111

```

```

Query Match          100.0%; Score 1014; DB 13; Length 1421;
Best Local Similarity 100.0%; Pred. No. 1.2e-281;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      217 ATGAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276

Qy      61 CTGAAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      277 CTGAAAGACAATTTGGCCATTTTGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 336

Qy      121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      337 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 396

Qy      181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 456

Qy      241 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      457 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 516

Qy      301 TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCC'ACTGTGGAGTATATTAGT 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      517 TTAAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCC'ACTGTGCAGTATATTAGT 576

Qy      361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
        |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636

```


QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
 |||
 Db 637 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 696
 |||
 QY 481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||
 Db 697 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 756
 |||
 QY 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 |||
 Db 757 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 816
 |||
 QY 501 TTCTTAGAACAAAAATACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 |||
 Db 817 TTCTTAGAACAAAAATACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 876
 |||
 QY 561 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 |||
 Db 677 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 936
 |||
 QY 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
 |||
 Db 937 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 996
 |||
 QY 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 |||
 Db 881 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 1056
 |||
 QY 841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 900
 |||
 Db 1057 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 1116
 |||
 QY 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 |||
 Db 1117 CTCATTGAGTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1176
 |||
 QY 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 |||
 Db 1177 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1230
 |||

RESULT 2

US-10-037-270-111

; Sequence 111, Application US/10037270

; Publication No. US20030104529A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yunqing
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Tillinghast, John
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/10/037,270
 ; CURRENT FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1104
 ; SOFTWARE: pt_FL_genes Version 1.0
 ; SEQ ID NO 111
 ; LENGTH: 1421
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (217)..(1230)
 03-10-037-270-111

Query Match 100.0%; Score 1014; DB 15; Length 1421;
 Basic Local Similarity 100.0%; Pred. No. 1.2e-281;
 Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAAAAATGCCTTTGTTTGTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 60
 |||
 DB 117 ATGAAAAAAAAATGCCTTTGTTTGTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAATC 276
 QY 61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 120
 |||
 DB 277 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAAGACAGACAAGGCTTCAGAA 336
 QY 121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
 |||
 DB 337 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 396
 QY 181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
 |||
 DB 397 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 456
 QY 241 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
 |||
 DB 457 AACTGATAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 516
 QY 301 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 360
 |||
 DB 517 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGT 576
 QY 361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
 |||

Db 577 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 636
 QY 421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
 |||||
 Db 637 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 696
 QY 421 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
 |||||
 Db 697 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 756
 QY 541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
 |||||
 Db 757 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 816
 QY 601 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
 |||||
 Db 817 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 876
 QY 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 |||||
 Db 877 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 936
 QY 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
 |||||
 Db 937 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 996
 QY 781 CTCCTTCGGGATAAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 840
 |||||
 Db 997 CTCCTTCGGGATAAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTT 1056
 QY 841 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 900
 |||||
 Db 1057 GTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCAA 1116
 QY 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 |||||
 Db 1117 CTCATTGAGTTTCTGAGCAGCTTCCAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1176
 QY 961 GAGAAGAAGTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 |||||
 Db 1177 GAGAAGAAGTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1230

RESULT 3

US-10-025-730-2

; Sequence 2, Application US/10025730

; Publication No. US20030045466A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina A.

; TITLE OF INVENTION: CALCIUM BINDING PROTEIN

; FILE REFERENCE: PF-0635 US

; CURRENT APPLICATION NUMBER: US/10/025,730

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: US/09/190,965

; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 3734805
US-10-025-730-2

Query Match 99.7%; Score 1010.8; DB 15; Length 1344;
Best Local Similarity 99.8%; Pred. No. 1e-280;
Matches 1012; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY      1 ATGAAAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 60
      |||
Db      124 ATGAAAAAAAAATGCCTTTGTTTAGTAAATCACACAAAAATCCAGCAGAAATTGTGAAAAATC 183

QY      61 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 120
      |||
Db      184 CTGAAAGACAATTTGGCCATTTTGGAAAAGCAAGACAAAAGACAGACAAGGCTTCAGAA 243

QY     121 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 180
      |||
Db     244 GAAGTGTCTAAATCACTGCAAGCAATGAAAGAAATTCTGTGTGGTACAAACGAGAAAGAA 303

QY     181 CCCCCAACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTAGTG 240
      |||
Db     304 CCCCCGACAGAAGCAGTGGCTCAGCTAGCACAAAGAACTCTACAGCAGTGGCCTGCTGCTG 363

QY     361 AACTGTAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 300
      |||
Db     364 AACTGTAGCTGACCTGCAGCTGATAGACTTTGAGGGAAAAAAGATGTGACCCAGATA 423

QY     361 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCTACTGTGGAGTATATTAGT 360
      |||
Db     424 TTTAACAACATCTTGAGAAGACAGATAGGCACTCGGAGTCTACTGTGGAGTATATTAGT 483

QY     361 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 420
      |||
Db     424 GCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTA 543

QY     421 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 480
      |||
Db     544 CGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACTTGCCAAAATCATCCTC 603

QY     481 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 540
      |||
Db     604 TTTTCTAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCT 663

QY     541 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 600
      |||
Db     664 TCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATAAAGTGTGGTAGCAGAC 723

QY     601 TTCTTAGAACAAAATTACGACACTATTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 660
      |||
```


Db 724 TTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAG 783

Qy 661 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 720
 |||

Db 734 AATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCAC 843

Qy 721 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 780
 |||

Db 844 AACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAAC 903

Qy 781 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTAAGGTGTTT 840
 |||

Db 904 CTCCTTCGGGATAAAAGTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTAAGGTGTTT 963

Qy 841 GTGCCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 900
 |||

Db 964 GTGCCCAGTCCTCACAAAACACAGCCTATTGTGGAGATCCTGTTAAAAAATCAGCCCCAA 1023

Qy 901 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 960
 |||

Db 1024 CTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGAC 1083

Qy 961 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1014
 |||

Db 1084 GAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAGAAAACGGCCCCCTTGA 1137

RESULT 4

US-09-918-995-5343

Sequence 5343, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO 5343

LENGTH: 475

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(475)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-5343

Query Match 39.3%; Score 398; DB 11; Length 475;

Best Local Similarity 100.0%; Pred. No. 3.5e-104;

Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 617 ACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGA 676


```

Db      1 ACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGA 60
Qy      677 GACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTTGCCATCATGA 736
Db      61 GACAGTCTTTAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTTGCCATCATGA 120
Qy      737 CAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAA 796
Db      121 CAAAGTATATCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAA 180
Qy      797 GTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACA 856
Db      181 GTCCCAACATCCAGTTTGAAGCCTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACA 240
Qy      857 AAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGA 916
Db      241 AAACACAGCCTATTGTGGAGATCCTGTAAAAAATCAGCCCAAACCTCATTGAGTTTCTGA 300
Qy      917 CCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGA 976
Db      301 GCAGCTTCCAAAAAGAAAGGACGGATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGA 360
Qy      977 TTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 1014
Db      361 TTAAACAGATCCGAGACTTGAAGAAAACGGCCCCTTGA 398

```

RESULTS

US-09-910-943-318

Sequence 318, Application US/09910943

Patent No. US20020081610A1

GENERAL INFORMATION:

APPLICANT: Hemmati-Brivanlou, Ali

APPLICANT: Altman, Curtis

TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression

FILE REFERENCE: 7529/1G148US1

CURRENT APPLICATION NUMBER: US/09/910,943

CURRENT FILING DATE: 2001-07-23

NUMBER OF SEQ ID NOS: 742

SOFTWARE: PatentIn version 3.1

SEQ ID NO 318

LENGTH: 690

TYPE: DNA

ORGANISM: Xenopus laevis

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(690)

OTHER INFORMATION: n may be a or g or c or t/u

US-09-910-943-318

Query Match 28.5%; Score 288.8; DB 9; Length 690;

Best Local Similarity 80.5%; Pred. No. 1.4e-72;

Matches 338; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```

Qy      595 GCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGAAATTGCTTCAG 654
          ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

Db      101 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACCTTCACAATTATCACAAAATACA 160
QY      746 TCAGCAAGCCGGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACA 805
Db      161 TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA 220
QY      806 TCCAGTTTGAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGC 865
Db      221 TCCAGTTTGAGGCCTTTTCACGTTTTTAAGGTGTTTGTAGCCAATCCTAACAAGACGCAGC 280
QY      866 CTATTGTCGAGATCCTGTATAAAAAATCAGCCCAAACCTCATTGAGTTTCTGAGCAGCTTCC 925
Db      231 CCATCCTAGACATCCTCCTCAAGAACCAGGCCAAACTCATAGAGTTCCTCAGCAAGTTTC 340
QY      926 AAAAAGAAAGCACGGATGATGAGCAGTTCGCTGACGAGAAGAAGTACTTGATTAAACAGA 985
Db      341 AGAACGACAGGACGGAGGATGAGCAGTTTAACGACGAGAAGACCTATTTAGTTAAACAGA 400
QY      986 TCCCAGACTTGAAGAAAACGGCCC 1009
Db      401 TCCGGGATTGAAGAGACCCGCTC 424

```

RESULT 7

US-09-867-701-5899

Sequence 5899, Application US/09867701

Patent No. US20020132237A1

GENERAL INFORMATION:

APPLICANT: Aglate, Paul A.

APPLICANT: Jones, Robert

APPLICANT: Harlocker, Susan L.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

FILE REFERENCE: 210121.497

CURRENT APPLICATION NUMBER: US/09/867,701

CURRENT FILING DATE: 2001-05-29

NUMBER OF SEQ ID NOS: 10912

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 5899

LENGTH: 447

TYPE: DNA

ORGANISM: Homo sapien

US-09-867-701-5899

Query Match: 24.1%; Score 244.8; DB 10; Length 447;
 Best Local Similarity 77.3%; Pred. No. 5.3e-60;
 Matches 297; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```

QY      626 TTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT 685
Db      41 TTTTCAGTGAATATGAGAAGTTACTTCATTGAGAAAATTATGTGACAAAAGACAGTCAC 100
QY      686 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATA 745
Db      101 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACCTTCACAATTATGACAAAATACA 160

```


[illegible]

US-09-867-701-4953

Patent No. US20020132237A1

APPLICANT: Aglate, Paul A.

APPLICANT: Harlocker, Susan L.

FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

FILE REFERENCE: 210121.497

CURRENT APPLICATION NUMBER: US/09/867,701

NUMBER OF SEQ ID NOS: 10912

; SEQ. ID NO 4953

LENGTH: 450

ORGANISM: Homo sapien

Query Match 24.1%; Score 244.8; DB 10; Length 450;

Matches 297; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy	626	TTTTGAAGACTATGAGAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTT	685
Db	27	TTTTCAGTGAATATGAGAAGTTACTTCATTTCAGAAAATTATGTGACAAAAGACAGTCAC	86
Qy	636	TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTGCCATCATGACAAAGTATA	745
Db	37	TGAAGCTTCTCGGTGAAC TACTACTAGATAGACACAACCTTCACAATTATGACAAAATACA	146
Qy	745	TCAGCAAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAAGTCCCAACA	805
Db	147	TCAGTAAACCTGAGAACCTCAAATTAATGATGAACCTGCTGCGAGACAAAAGTCGCAACA	206

27 TTTTCAGTGAATATGAGAAGTTACTTCATTTCAGAAAATTATGTGACAAAAGACAGTCAC 86

636 TAAAGCTGCTAGGGGAGCTGATCCTGGACCGTCACAACCTTGCCATCATGACAAAGTATA 745

37 TGAAGCTTCTCGGTGAACTACTACTAGATAGACACAACCTTCACAATTATGACAAAATACA 146

Qy. 745 TCAGCAAGCCGGAGAACCTGAACTCATGATGAACCTCCTTCGGGATAAAAAGTCCCAACA 805

[illegible]

147 TCAGTAAACCTGAGAACC TCAAA TAAATGATGAACC TGC TGC GAGACAAAAG TCCCAACA 200

Query Match 19.2%; Score 195; DB 10; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.1e-45;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      820 TTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 879
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      387 TTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 328

QY      880 CTGTTAAAAAATCAGCCCAAACCTATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACG 939
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      327 CTGTTAAAAAATCAGCCCAAACCTATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACG 268

QY      940 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 999
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      267 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 208

QY      1000 AAAACGGCCCCCTTGA 1014
      ||||||||||||
Db      207 AAAACGGCCCCCTTGA 193
```

RESULT 11

US-09-330-107-481/c

; Sequence 481, Application US/09880107

; Patent No. US20020142981A1

GENERAL INFORMATION:

; APPLICANT: Horne, Darci T.

; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherf, Uwe

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-WO

; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379

; PRIOR FILING DATE: 2000-06-14

; PRIOR APPLICATION NUMBER: US 60/237,054

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 3950

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 481

; LENGTH: 387

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA234362

US-09-330-107-481

Query Match 19.2%; Score 195; DB 10; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.1e-45;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      820 TTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 879
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      387 TTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATTGTGGAGATC 328

QY      880 CTGTTAAAAAATCAGCCCAAACCTATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACG 939
```



```

Db      327 CTGTTAAAAAATCAGCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAAGAAAGGACG 268
QY      940 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 999
Db      267 GATGATGAGCAGTTCGCTGACGAGAAGAACTACTTGATTAAACAGATCCGAGACTTGAAG 208
QY      1000 AAAACGGCCCCTTGA 1014
Db      207 AAAACGGCCCCTTGA 193

```

RESULT 12

US-10-257-826A-113

```

; Sequence 113, Application US/10257826A
; Publication No. US20030181407A1
; GENERAL INFORMATION:
; APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
; APPLICANT: PALIN, Marie-France
; APPLICANT: POMAR, Candido
; APPLICANT: GARIEPY, Claude
; TITLE OF INVENTION: Steatosis-modulating factors and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 14654-2US
; CURRENT APPLICATION NUMBER: US/10/257,826A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 60/197936
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: PCT/CA01/00509
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial sequence
; OTHER INFORMATION: Muscular steatosis
; OTHER INFORMATION: Porcine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(722)
; OTHER INFORMATION: n = A,T,C or G
US-10-257-826A-113

```

```

Query Match      16.7%; Score 169.8; DB 13; Length 722;
Best Local Similarity 60.1%; Pred. No. 3.1e-38;
Matches 303; Conservative 0; Mismatches 196; Indels 5; Gaps 4;

```

```

QY      347 TGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCC 406
Db      2   TGGTGAATNCCTCTGCCCCACNGAATTTTTGGTCATGGTANTNGAAGGGGATNAATNTT 67
QY      407 CACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTTCGACATGAACCACTTG 466
Db      68  CCGAAATTNCNNTTAATTGGGGNATNATGGTNAGANAATGCCTTNGACCTNNACCGCTTG 127

```


QY 467 CAAAATCATCCTCTTTTC--TAATCAATTCAGAGATTTCTTTAAGTACGTGGAGTTGTC 524
 |||||
 Db 128 CAAAATCATTTTGNNGGCCGAACACAGTTTATAGAGATCTTCACATATGTCTAAATGTN 187
 QY 525 AACA-TTGTATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTTACTAACCAGACATA 583
 |||||
 Db 138 ANCATTTTACATATCTTTACATNCNNTTNCNCATTTTNNGNNTTACTTTCACGACATA 247
 QY 584 AAGTGTGGTAGCAGACTTCTTAGAACAAAATTACGACACTATTTTTGAAGACTATGAGA 643
 |||||
 Db 248 TATTGCTCACNGCGCAANTTTTGAACANCATTATGATANATTTTTCAGTGAATATGATG 307
 QY 644 AATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGAGC 703
 |||||
 Db 308 AAGNGCTTCATTCTTAAATTATGTGGCCACAAGACAATCACTGAAGCTTCTCGGNGAAC 367
 QY 704 TGATCCTGGACCGTCACAACCTTTGCCATCATGACAAAGTATATCAGCAAGCCGGAGAACC 763
 |||||
 Db 368 TACTACTANATAGACNCNACTTCNCANTATGACCACATACCTCATTAAACCTGNGNACC 427
 QY 764 TGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAA-CATCCAGTTTGAAGCCTTT 822
 |||||
 Db 428 T-CCATTAATGATGAACCTGCCTGCAGACAAAAGTCGGAACCTTCCANTTTGAGGGCTTN 486
 QY 822 CATGTTTTTAAGGTGTTTGTGGCC 846
 |||||
 Db 487 CACGTTTTTAANGGGNTGTNNC 510

SEQUENCE 119

US-10-257-326A-119

; Sequence 119, Application US/10257826A
 ; Publication No. US20030181407A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
 ; APPLICANT: PALIN, Marie-France
 ; APPLICANT: POMAR, Candido
 ; APPLICANT: GARIEPY, Claude
 ; TITLE OF INVENTION: Steatosis-modulating factors and uses
 ; TITLE OF INVENTION: thereof
 ; FILE REFERENCE: 14654-2US
 ; CURRENT APPLICATION NUMBER: US/10/257,326A
 ; CURRENT FILING DATE: 2002-10-17
 ; PRIOR APPLICATION NUMBER: 60/197936
 ; PRIOR FILING DATE: 2000-04-17
 ; PRIOR APPLICATION NUMBER: PCT/CA01/00509
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 305
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 119
 ; LENGTH: 700
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Artificial sequence
 ; OTHER INFORMATION: Muscular steatosis


```
; OTHER INFORMATION: Porcine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(700)
; OTHER INFORMATION: n = A,T,C or G
US-10-257-826A-119
```

```
Query Match          16.4%; Score 166.6; DB 13; Length 700;
Best Local Similarity 60.1%; Pred. No. 2.6e-37;
Matches 304; Conservative 0; Mismatches 197; Indels 5; Gaps 4;
```

```
QY      345 TGTGGAGTATATTAGTGCTCATCCTCATATCCTGTTTATGCTCCTCAAAGGATATGAAGC 404
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      5 TCTGGTGAATCCCTCTGCCCCACNGAATTTTGGTCATGGTANTNGAAGGGGATNAATN 65

QY      405 CCCACAGATTGCCTTACGTTGTGGGATTATGCTGAGAGAATGTATTCGACATGAACCACT 464
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      66 TTCCGAAATTTTCGATTAATTGGGGNATNATGGTNAGANAATGCCTTNGACCTCCACCGCT 125

QY      465 TGCCAAAATCATCCTCTTTTC--TAATCAATTCAGAGATTTCCTTAAAGTACCTGGAGTTG 522
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      126 TGCCAAAATCATTTTGNNGGCCGAACACAGTTTATAGAGATCTTCACATATGTCTAAATG 185

QY      583 TCAACA-TTTGATATTGCTTCAGATGCCTTTGCTACTTTCAAGGATTACTAACCAGACA 581
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      136 TNCACATTTTACATATCTTTACATNCNNTTNCNCATTTTNNGNNTTACTTTTCACGACA 245

QY      642 TAAAGTGTGTGGTAGCAACTTCTTGAACAAAATTACGACACTATTTTGAAGACTATGA 641
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      246 TATATTGCTCACNGCGCAANTTTTGAACANCATTATGATANATTTTTCAGTGAATATGA 305

QY      702 GAAATTGCTTCAGTCTGAGAATTATGTTACTAAGAGACAGTCTTTAAAGCTGCTAGGGGA 701
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      306 TGAAGNGCTTCATTCTTAAAATTATGTGGCCACAAGACAATCACTGAAGCTTCTCGGNGA 365

QY      762 GCTGATCCTGGACCGTCACAACTTTGCCATCATGACAAAGTATATCAGCAAGCCGGGAGAA 761
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      366 ACTACTACTANATAGACNCNACTTCNCCANTATGACCACATACCTCATTAAACCTGNGNA 425

QY      820 CCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAA-CATCCAGTTTGAAGCCT 820
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      426 CCT-CCATTAATGATGAACCTGCCTGCAGACAAAAGTCGGAACCTTCCANTTTGAGGGCT 484

QY      821 TTCATCTTTTAAAGGTGTTTGTGGCC 846
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      485 TNCACGTTTTTAAANGGGGNTGTNNNC 510
```

RESULT 14

```
US-09-770-445-592
; Sequence 592, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
```


QY 751 AAGCCGGAGAACCTGAAACTCATGATGAACCTCCTTCGGGATAAAAGTCCCAACATCCAG 810
 | | | | | | | | | | | | | | | | | | | | | |
 Db 372 TCGATGGATAACTTGAGGATTCTGATGAATCTTCTCAGAGAATCAAGCAAGACTATTTCAG 431
 QY 811 TTTGAAAGCCTTTTCATGTTTTTAAGGTGTTTGTGGCCAGTCCTCACAAAACACAGCCTATT 870
 | | | | | | | | | | | | | | | | | | | | | |
 Db 432 ATAGAAGCTTTCATGTTTTCAAGCTGTTTGTAGCGAACCAAAACAAGCCTTCAGACATC 491
 QY 871 GTGGAGATCCTGTATAAAAAATCAGCCCAAACTCATTGAGTTTCTGAGCAGCTTCCAAAAA 930
 | | | | | | | | | | | | | | | | | | | | | |
 Db 492 GCCAACATTCTGGTGGCAAACAGAAACAAGCTTCTGAGATTGTTGGCTGATATCAAGCCG 551
 QY 931 GAAAGGACGGATGATGAGCAGTTTCGCTGACGA 962
 | | | | | | | | | | | | | | | | | | | | | |
 Db 552 GACAAAGAGGACGAGAGGTTTGACGCAGACAA 583

RESULT 15

US-09-923-876-1251

; Sequence 1251, Application US/09923876

; Patent No. US20020013958A1

GENERAL INFORMATION:

; APPLICANT: Lalgudi, Raghunath V.

; APPLICANT: Kamigaki, Laura Y. (Ito)

; APPLICANT: Sherman, Bradley K.

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN

SEQUENCING

; FILE REFERENCE: PL-0012-1 CON

; CURRENT APPLICATION NUMBER: US/09/923,876

; CURRENT FILING DATE: 2001-08-06

; PRIOR APPLICATION NUMBER: 09/298,329

; PRIOR FILING DATE: 1999-04-21

; PRIOR APPLICATION NUMBER: 60/085,331

; PRIOR FILING DATE: 1998-05-05

; NUMBER OF SEQ ID NOS: 6332

; SOFTWARE: PERL Program

; SEQ ID NO 1251

; LENGTH: 262

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No. US20020013958A1 7C0158378H1

; NAME/KEY: unsure

; LOCATION: 148

; OTHER INFORMATION: a, t, c, g, or other

US-09-923-876-1251

Query Match 7.3%; Score 74.2; DB 9; Length 262;
 Best Local Similarity 55.5%; Pred. No. 6.4e-11;
 Matches 142; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 311 TCTTGAGAAGACAGATAGGCACTCGGAGTCCTACTGTGGAGTATATTAGTGCTCATCCTC 370
 | | | | | | | | | | | | | | | | | | | | | |
 Db 7 TCTTGAGACAGAAGGTTGATGAAAGCTATTGTTGCGTCCAGTATATTGAAAATCATTTTG 66

QY 371 ATATCCTGTTTATGCTCCTCAAAGGATATGAAGCCCCACAGATTGCCTTACGTTGTGGGA 430
 || | || | | | | | | || | | | | | | |
 Db 57 ATCTTTTGGATTTCCTTGTTGTTTGCTATAAGAACTTGGAAGTCGCGTTGAATTGTGGAA 126
 QY 431 TTATGCTGAGAGAATGTATTTCGACATGAACCACTTGCCAAAATCATCCTCTTTTCTAATC 490
 || | || | || | || | | | | | | | | | | | | |
 Db 127 ACATGTTGCGAGAATGCATAANATATCCTACACTTGCAAAATATATATTGGAGTCAAGCA 186
 QY 491 AATTCAGAGATTTCTTTAAGTACGTGGAGTTGTCAACATTTGATATTGCTTCAGATGCCT 550
 || | || | || | || | | | | | | | | | | | | |
 Db 187 GCTTCGAGTTGTTTTTCCAGTATGTTGAATTGTCAAACCTTCGATATTGCATCTGATGCTC 246
 QY 551 TTGCTACTTTTCAAGGA 566
 | || | || | || |
 Db 247 TGAACACTTTTCAAGGA 262

Search completed: January 6, 2004, 05:04:45
Job time : 1400 secs